(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
11 March 2004 (11.03.2004)

PCT

(10) International Publication Number
WO 2004/020587 A2

(51) International Patent Classification[7]: C12N

(21) International Application Number:
PCT/US2003/026773

(22) International Filing Date: 25 August 2003 (25.08.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/406,341 28 August 2002 (28.08.2002) US

(71) Applicant *(for all designated States except US)*: **SCHERING AKTIENGESELLSCHAFT** [DE/DE]; Mueller-strasse 178, 13342 Berlin (DE).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **McCARRICK, Margaret, Anne** [US/US]; 3721 Balboa Terrace, Unit E, San Diego, CA 92117 (US). **MORSER, Michael, John** [GB/US]; 746 Sanchez Street, San Francisco, CA 94114 (US). **NAGASHIMA, Mariko** [US/US]; 2409 Palmer Avenue, Belmont, CA 94002 (CA). **ZHAO, Lei** [CN/US]; 21 Palmer Court, Pleasant Hill, CA 94523 (US).

(74) Agent: **WASHTIEN, Wendy, L.**; Berlex Biosciences, 2600 Hilltop Drive, P.O. Box 4099, Richmond, CA 94804 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 2004/020587 A2

(54) Title: CARBOXYPEPTIDASE B RELATED POLYPEPTIDES AND METHODS OF USE

(57) Abstract: Recombinant TAFI polypeptide mutants, as well as the polynucleotides encoding such polypeptides, are disclosed. Also disclosed are methods for utilizing such polypeptides in methods of screening for potential therapeutic agents, substrates, inhibitors or enhancers of carboxypeptidases.

CARBOXYPEPTIDASE B RELATED POLYPEPTIDES AND METHODS OF USE

This application claims the benefit of U.S. Provisional Application Serial No. 60/406,341, filed on August 28, 2002, which is incorporated herein by reference.

FIELD OF THE INVENTION

The present invention relates to thrombin-activatable fibrinolysis inhibitor (TAFI) polypeptide mutants and polynucleotides encoding such polypeptide mutants. It also relates to the use of such polypeptide mutants in screening for inhibitors and/or substrates and/or enhancers of TAFI, carboxypeptidase N and other related carboxypeptidases.

BACKGROUND OF THE INVENTION

Carboxypeptidases (CPs) are enzymes that catalyze the hydrolysis of the C-terminal peptide bond in peptides and proteins. The removal of one or a few amino acids from the C-terminus of a peptide or protein can have profound effects on the biological activity of that molecule. CPs are involved in many diverse processes in the body, including digestion and assimilation of dietary proteins, processing of peptide hormone precusors, regulation of peptide hormone activity and regulation of protein binding. (Skidgel and Erdos, *Handbook of Proteolytic Enzymes*, Barret et al. (eds), pp. 1344-1349, (1998) Academic Press, San Diego).

Thrombin Activatable-Fibrinolysis Inhibitor (TAFI [EC 3.4.17.20]), which is also known as plasma carboxypeptidase B (PCPB), carboxypeptidase U, and carboxypeptidase R, is a plasma protein that has basic carboxypeptidase activity. TAFI (SEQ ID NO: 2) is synthesized as a 423 amino acid polypeptide Following cleavage of a 22 amino acid signal sequence, the polypeptide is secreted and then activated by the thrombin/thrombomodulin complex to produce activated TAFI (TAFIa; SEQ ID NO: 3).

TAFIa is the active carboxypeptidase and cleaves C-terminal basic residues that are newly exposed by the cleavage of fibrin by plasmin (Nesheim et al., *Thromb. Haemost.*, 78:386-91 (1997); Bouma et al., *Thromb. Res.*, 101:329-54 (2001); Schatteman et al., *Clin. Appl. Thromb. Hemost.*, 7:93-101 (2001)). Since these C-terminal basic residues are high affinity binding sites for both plasminogen and plasminogen activator, they serve as an amplification system for plasmin production. Activated TAFI removes these C-terminal basic residues, thereby reducing plasmin production, which leads to slower lysis of the clot. Activated TAFI may also be involved in plasmin-mediated cell migration (Plow et al., *Trends Cardiovasc. Med.*, 7:71-5 (1997)). *In vitro,* activated TAFI can also cleave a number of peptides which are biologically active in the circulation such as bradykinin and anaphylatoxins (Tan et al., *Biochemistry*, 34:5611-5816 (1995); Campbell et al., *Immunological Reviews*, 180:162-7 (2001)). *In vivo,* however, these peptides are known to be degraded rapidly by the second carboxypeptidase present in plasma, carboxypeptidase N (CPN, [EC 3.4.17.3]).

Both TAFI and CPN belong to a class of metallocarboxypeptidases that catalyze the hydrolysis of the C-terminal peptide bond in peptides and proteins. Based on sequence analysis, these metallocarboxypeptidases can be divided into two groups: 1) carboxypeptidases A1, A2, B, TAFI and mast cell carboxypeptidase A, and 2) carboxypeptidases N, H/E, M, D and Z (Aviles et al., *Eur J Biochem.*, 211, 381-9 (1993); Skidgel R.A. *Zinc Metalloproteases in Health and Diseases*. (Hooper, N. M. ed), Taylor and Francis Ltd., London, pp241-283 (1996); Reznik et al., *Cell Mol. Life Sci.*, 58:1790-1804 (2001)). The sequence similarity is high within each group (40-58%), but much lower between the groups (14-20%).

Carboxypeptidases from the first group are synthesized as inactive zymogens and require removal of a propeptide segment before they exert carboxypeptidase activity, optimally at neutral pH (i.e. TAFI to TAFIa). In their active form, these CPs have a preference for aromatic, aliphatic, or basic residues and their activity is inhibited by a naturally occurring small protein, carboxypeptidase inhibitor from potato (CPI), as well as by synthetic compounds such as guanidinoethylmercaptosuccinic acid (GEMSA). The three-dimensional structure and the mechanism of action for pancreatic carboxypeptidases A and B have been widely studied, and are found to be very similar.

On the other hand, carboxypeptidases from the second group (such as CPN) are constitutively active towards basic residues with various pH optimums, and possess a long C-terminal extension whose function is unknown. This second group of carboxypeptidases is inhibited by GEMSA, but they are not susceptible to inhibition by CPI. This group of carboxypeptidases perform important functions in regulation of biologically active peptides such as processing of propeptide hormones, inactivation of active peptides and alteration of substrate specificity for receptor binding (Steiner et al., *J. Biol. Chem.* 267:23435-23438 (1992); Bhoola et al., *Pharmacol. Rev.* 44:1-80 (1992); Bokisch and Muller-Eberhard, *J. Clin. Invest.* 49:2427-2436 (1970)).

Recently, the crystal structure of CPD was elucidated and reveals an overall topological similarity to that of CPA and CPB with some insertions and deletions, but with unique structural features that may explain the differences in the activities of the two groups of CPs (i.e. substrate specificity, inhibitor susceptibility) (Gomis-Ruth et al., *EMBO*, 18:5817-26 (1999)).

One can take advantage of the similarities and differences between the two groups of carboxypeptidases to create polypeptide mutants useful for screening and design of substrates/inhibitors/enhancers of these enzymes.

SUMMARY OF THE INVENTION

The present invention provides isolated thrombin-activatable fibrinolysis inhibitor (TAFI) polypeptide mutants, polynucleotides encoding such polypeptides, and variants thereof, useful for screening inhibitors/modulators of carboxypeptidase activity. In particular, the invention provides TAFI polypeptide mutants having one or more mutations at an amino acid residue within the sequence of wild-type TAFI (SEQ ID NO: 2), wherein said residue is involved in substrate binding.

Preferably, the TAFI polypeptide mutant has a mutation in amino acid position 207, 248, 256, or 257, wherein the numbering of the amino acid positions is taken from the amino acid sequence of activated TAFI (SEQ ID NO: 3). Particularly preferred embodiments are TAFI polypeptide mutants in which: the Ser at amino acid position 207 is replaced with an Asp; the Leu at amino acid position 248 is replaced with a Tyr; the Asp at amino acid position 256 is replaced with a Glu; or the Asp at amino acid position 257 is replaced with an Ala.

In another preferred embodiment, the TAFI polypeptide mutant is one in which both the Ser at amino acid position 207 has been replaced with an Asp and the Asp at amino acid position 256 has been replaced with a Glu.

In a related aspect, the present invention further provides for TAFI polypeptide mutants having at least one mutation, preferably at position 207, 248, 256, or 257, and a second mutation at an amino acid position selected from positions 69, 125, 143, 197, 198, 199, 249, or 271.

In another aspect, the present invention provides a polynucleotide encoding a TAFI polypeptide mutant of the present invention.

In a related aspect, the invention provides a recombinant vector having a polynucleotide encoding a TAFI polypeptide mutant, where the polynucleotide is operably linked to at least one regulatory sequence such that the encoded polypeptide is expressed in cells.

In another aspect, the present invention provides a method of screening for agents which modulate the activity of a TAFI polypeptide mutant of the invention or a related carboxypeptidase, wherein the method comprises: 1) contacting a cell expressing said TAFI polypeptide mutant or a tissue cell extract thereof with a putative modulatory agent and 2) measuring the activity of said polypeptide.

In another aspect, the present invention provides a method of screening for agents which modulate the expression of a polynucleotide encoding a TAFI polypeptide mutant of the invention or a related carboxypeptidase, where the method comprises: 1) contacting a cell expressing said TAFI polypeptide mutant or a tissue cell extract thereof with a putative modulatory agent and 2) measuring the amount or activity of said polypeptide.

In another aspect, the present invention provides a method for altering the binding specificity of the polypeptide of SEQ ID NO: 2 or a related carboxypeptidase, comprising mutating or altering at least one amino acid residue within the polypeptide sequence involved in substrate binding.

In another aspect, the present invention provides a method of screening for an agent which binds to a TAFI polypeptide mutant of the invention, comprising 1) contacting the isolated polypeptide with a putative binding agent and 2) determining the presence of a complex between the agent and the polypeptide.

BRIEF DESCRIPTION OF THE DRAWINGS

The foregoing and other objects of the present invention, the various features thereof, as well as the invention itself may be more fully understood from the following description, when read together

with the accompanying drawings.

Figure 1: Amino acid sequence alignment between TAFIa and Carboxypeptidase N. The alignment was done manually using SeqLab based on the results from the secondary structure prediction program Psi-Pred. Numbering for TAFIa (SEQ ID NO: 3) starts with the alanine following the site of cleavage of the activation peptide. This corresponds to amino acid residue 115 of wild-type TAFI (SEQ ID NO: 2). Residues that are known or predicted to be involved in zinc binding, catalytic activity and substrate binding are indicated.

Figure 2: Activation of wild type TAFI and TAFI polypeptide mutants by the thrombin/thrombomodulin complex. Ten μl each of non-activated and activated wild type TAFI and TAFI polypeptide mutants were run on 10% SDS/PAGE under denatured conditions, followed by Western blot using polyclonal Sheep-antiTAFI-Ig. Activation was performed as described in "Example III". Samples are non-activated and activated wild type TAFI (lanes 1 and 2), non-activated and activated D256Q (lanes 3 and 4), non-activated and activated D257A (lanes 5 and 6), non-activated and activated D256Q/ S207D (lanes 7 and 8), ono-activated and activated S207D (lanes 9 and 10), and non-activated and activated L248W (lanes 11 and 12).

Figure 3: Carboxypeptidase B-like activity of activated wild type TAFI, TAFI polypeptide mutants, CPN and CPA. Wild type TAFI and TAFI polypeptide mutants were first activated as described in "Example III". The activity measurements for the activated wild type TAFI, the TAFI polypeptide mutants, CPN, and carboxypeptidase A (CPA) using TAFI developer were performed as described in "Example III". The activity data were expressed as percentage of the activity of wild type TAFI. Each column represents the average of the duplicates of two experiments.

Figure 4: CPI inhibition of the activity of activated wild type TAFI and TAFI polypeptide mutants. Wild type TAFI and TAFI polypeptide mutants were first activated as described in "Example III", and the carboxypeptidase activities were measured with TAFI developer in the presence of CPI at various concentrations. The percentage inhibition of carboxypeptidase activity was calculated for each CPI concentration.

Figure 5: Comparison of carboxypeptidase activities of activated wild type TAFI, TAFI polypeptide mutants, CPN, and CPA. Wild type TAFI and TAFI polypeptide mutants were first activated as described in "Example III". The activity measurements for the activated wild type TAFI, the TAFI polypeptide mutants, CPN, and CPA using different hippuryl-amino acid substrates were performed as described in "Example III". The activity data were **normalized** to percentage of the activity of wild type TAFI. Each column represents the average of the duplicates of two experiments

Figures 6A and 6B: The effect of wild type TAFI and TAFI polypeptide mutant s on human plasma clot lysis. 6A) Wild type TAFI without and with CPI (open circle, and open square), D256Q without and with CPI (open diamond and x), S207D without and with CPI (+ and open triangle), L248W without and with CPI (closed circle and closed square), D256Q/S207D without and with CPI (closed diamond and closed triangle facing up), and D257A without and with CPI (closed triangle facing down and open circle with a dot). 6B) Samples were added to TAFI deficient human plasma with thrombin, calcium, tissue-type plasminogen activator and thrombomodulin as described in "Example IV". Clot formation and lysis was monitored by measuring absorbance at 405 nm. Figure legend: No TAFI, no CPI (open circle), wild type TAFI without and with CPI (open square and open diamond).

Figure 7: Superposition of the side-chains from TAFI and CPN. The human TAFI model was built based on the human CPB structure (PDB1KWM, chain A, not shown) in complex with benzamidine). The human CPN model was built based on the duck CPD structure (PDB1QMU, not shown).

DETAILED DESCRIPTION OF THE INVENTION

Definitions

Technical and scientific terms used herein have the meanings commonly understood by one of ordinary skill in the art to which the present invention pertains, unless otherwise defined. Reference is made herein to various methodologies known to those of ordinary skill in the art. Publications and other materials setting forth such known methodologies to which reference is made are incorporated herein by reference in their entireties as though set forth in full. Standard reference works setting forth the general principles of recombinant DNA technology include Sambrook, J., *et al.* (1989) *Molecular Cloning,: A Laboratory Manual,* 2d Ed., Cold Spring Harbor Laboratory Press, Planview, N. Y.; McPherson, M. J., Ed. (1991) *Directed Mutagenesis: A Practical Approach,* IRL Press, Oxford; Jones, J. (1992) *Amino Acid and Peptide Synthesis, Oxford Science Publications*, Oxford; Austen, B. M. and Westwood, O. M. R. (1991) *Protein Targeting and Secretion*, IRL Press, Oxford. Any suitable materials and/or methods known to those of ordinary skill in the art can be utilized in carrying out the present invention; however, preferred materials and/or methods are described. Materials, reagents and the like to which reference are made in the following description and examples are obtainable from commercial sources, unless otherwise noted.

As used herein, "polypeptide" refers to a full-length protein or fragment thereof, or peptide.

As used herein, "wild type" refers to a nucleic acid or polypeptide molecule having the same nucleotide and/or amino acid sequence as a naturally-occurring molecule.

As used herein, "TAFI" refers to the TAFI polypeptide polymorph (SEQ ID NO: 2) as described by (Zhao et al., *Thromb. Haemost.*, 80:949-55 (1998))., which is processed by removal of a 22 amino acid signal sequence. The term "TAFI zymogen" refers to the TAFI molecule once

the signal sequence has been cleaved. The TAFI zymogen is further cleaved to produce "activated TAFI" which is referred to as "TAFIa" and has the sequence SEQ ID NO: 3.

As used herein, a "TAFI polypeptide mutant" or grammatical equivalents thereof (e.g., TAFI mutant, mutant TAFI, TAFI mutant polypeptide, mutant TAFI polypeptide) refers to an TAFI polypeptide, or variant thereof, having at least one mutation in an amino acid residue corresponding to a position involved in substrate binding, in which the numbering of the residues which have been mutated is based on the numbering for activated TAFI (i.e. TAFIa; SEQ ID NO: 3)

As used herein, "variant" with reference to a polypeptide or polynucleotide, refers to a polypeptide or polynucleotide that may vary in primary, secondary, or tertiary structure, as compared to a reference polypeptide or polynucleotide, respectively (*e.g.*, as compared to a wild-type polypeptide or polynucleotide). For example, the amino acid or nucleic acid sequence may contain a mutation or modification that differs from a reference amino acid or nucleic acid sequence. In some embodiments, an TAFI variant may be a different isoform or polymorphism. Variants can be naturally-occurring, synthetic, recombinant, or chemically modified polypeptides or polynucleotides isolated or generated using methods well known in the art.

As used herein, "mutation" with reference to a polypeptide or polynucleotide, refers to a naturally-occurring, synthetic, recombinant, or chemical change or difference to the primary, secondary, or tertiary structure of a polypeptide or polynucleotide, as compared to a reference polypeptide or polynucleotide, respectively (*e.g.*, as compared to a wild-type polypeptide or polynucleotide). Polypeptides and polynucleotides having such mutations can be isolated or generated using methods well known in the art.

As used herein, the term "activity" refers to a molecule having a function or action. Activity includes enzymatic activity, wherein the molecule is an enzyme e.g., a carboxypeptidase, that recognizes, binds, cleaves and/or modifies a substrate.

As used herein, the term "modified" refers to a molecule with an amino acid or nucleotide sequence differing from a naturally-occurring, i.e., wild-type amino acid or nucleotide, sequence. A modified molecule can retain the function or activity of a wild-type molecule.

As used herein, a first nucleotide or amino acid sequence is said to be "similar" to a second sequence when a comparison of the two sequences shows that they have few sequence differences (i.e., the first and second sequences are nearly identical). For example, two sequences are considered to be similar to each other when the percentage of nucleotides or amino acids that differ between the two sequences can be between about 60% to 99.99%.

As used herein, the term "complementary" refers to nucleic acid molecules having purine and pyrimidine nucleotide bases which have the capacity to associate through hydrogen bonding to form base pairs thereby mediating formation of double stranded nucleic acid molecules. The following base pairs are related by complementarity: guanine and cytosine; adenine and thymine; and adenine and uracil. Complementary applies to all base pairs comprising two single-stranded nucleic acid

molecules, or to all base pairs comprising a single-stranded nucleic acid molecule folded upon itself.

As used herein, the term "conservative" refers to substitution of an amino acid residue for a different amino acid residue that has similar chemical properties. A conservative amino acid substitution includes: substituting any hydrophobic (e.g., nonpolar) amino acid for any other hydrophobic amino acid; or substituting any hydrophilic (polar, uncharged) amino acid for any other hydrophilic amino acid; or substituting any positively charged amino acid for any other positively charge amino acid; or substituting any negatively charge amino acid for any other negatively charged amino acid (TE Creighton, "Proteins" WH Freeman and Company, New York). The amino acid substitutions include, but are not limited to, substituting any of isoleucine (I), valine (V), and leucine (L) for any other of these hydrophobic amino acids; aspartic acid (D) for glutamic acid (E) and vice versa; glutamine (Q) for asparagine (N) and vice versa; and serine (S) for threonine (T) and vice versa. Other substitutions can also be considered conservative, depending on the environment of the particular amino acid and its role in the three-dimensional structure of the protein. For example, glycine (G) and alanine (A), or glycine (G) and serine (S) can frequently be interchangeable, as can alanine (A) and valine (V). Methionine (M), which is relatively hydrophobic, can frequently be interchanged with leucine and isoleucine, and sometimes with valine. Lysine (K) and arginine (R) are frequently interchangeable in locations in which the significant feature of the amino acid residue is its charge and the differing pK's of these two amino acid residues are not significant. Still other changes can be considered conservative in particular environments.

As used herein, the term "nonconservative" refers to substituting an amino acid residue for a different amino acid residue that has different chemical properties. The nonconservative substitutions include, but are not limited to aspartic acid (D) being replaced with glycine (G); asparagine (N) being replaced with lysine (K); or alanine (A) being replaced with arginine (R).

The single-letter codes for amino acid residues include the following: A = alanine, R = arginine, N = asparagine, D = aspartic acid, C = cysteine, Q = Glutamine, E = Glutamic acid, G = glycine, H = histidine, I = isoleucine, L = leucine, K = lysine, M = methionine, F = phenylalanine, P = proline, S = serine, T = threonine, W = tryptophan, Y = tyrosine, V = valine.

As used herein, substitution of one amino acid for another within a wild-type TAFI polypeptide is represented by the single letter code for the original amino acid residue, followed by the number of the residue where the substitution has occurred, followed by the single letter code for the substituted amino acid; for example, D256Q, represents a modified TAFI polypeptide in which the aspartic acid at residue 256 of activated TAFIa (SEQ ID NO: 3) has been replaced with a glutamine.

As used herein, "involved in substrate binding" refers to amino acid residues which have been identified by three-dimensional modeling which are known to be, or are located in the vicinity of, amino acid residues critical for carboxypeptidase substrate binding
(Skidgel R.A. *Zinc Metalloproteases in Health and Diseases*. (Hooper, N. M. ed), Taylor and Francis Ltd., London, pp241-283 (1996).

As used herein, "modulate", with reference to TAFI or other carboxypeptidase activity, refers

to an increase, decrease, induction, or repression of such activity. In some embodiments, such increase, decrease, induction or rpession of carboxypeptidase activity is relative to a reference molecule, e.g. a TAFI wild-type or mutant polypeptide.

TAFI Polypeptide Mutants

The present invention provides TAFI polypeptide mutants, polynucleotides encoding such mutant polypeptides, and variants thereof, useful for screening modulators of carboxypeptidase activity. In particular, the invention provides TAFI polypeptide mutants having one or more mutations at an amino acid residue which is involved in substrate binding.

TAFI polypeptide mutants can be constructed by replacing specific amino acid residues of a member of the first group of carboxypeptidases (*i.e.*, carboxypeptidase A1, A2, B, TAFI and mast cell carboxypeptidase A) with residues from a member of the second group of carboxypeptidases (*i.e.*, carboxypeptidase N, H/E, M, D and Z) that are involved in substrate binding or the binding of inhibitors/enhancers or other agents (see Figure 7).

In one embodiment, the TAFI polypeptide mutants of the present invention have a mutation at amino acid residue 207, 248, 256, or 257, where the numbering of the amino acid residues is taken from the amino acid sequence of activated TAFI (i.e. TAFIa; SEQ ID NO: 3).

In a preferred embodiment, the TAFI polypeptide mutant has the Asp residue at position 256 replaced with a Gln residue (D256Q). In a second preferred embodiment, the TAFI polypeptide mutant has the Ser residue at position 207 replaced with an Asp residue (S207D). In another embodiment, the TAFI polypeptide has the Leu residue at position 248 replaced with a Trp residue (L248W). In another preferred embodiment, the TAFI polypeptide has the Asp residue at position 257 replaced with an Ala residue (D256Q). In another embodiment, the TAFI polypeptide mutant has the Asp at position 256 replaced with a Gln residue and the Ser residue at position 207 replaced with an Asp residue (D256Q, S207D).

In other embodiments, additional amino acid residues can be changed, individually or in combination with any of the amino acids changes described thus far (*e.g.,* D256Q, D257A, D256Q/S207D, S207D and L248W) in the activated TAFI amino acid sequence (SEQ ID NO:3): Possible amino acid positions where substitution could be made are amino acid residues 69, 125, 142, 143, 197, 198, 199, 249 and 271, which have been described as being critical to substrate-binding for various carboxypeptidases (Skidgel R.A. *Zinc Metalloproteases in Health and Diseases*. (Hooper, N. M. ed), Taylor and Francis Ltd., London, pp241-283 (1996); Barbosa-Pereira et al., *J. Mol. Biol.* 321:537-547 (2002)).

A TAFI polypeptide of the present invention may be a recombinant polypeptide, an isolated polypeptide or a synthetic or semi-synthetic polypeptide, or combinations thereof, preferably a recombinant polypeptide. As used herein, the terms polypeptide, oligopeptide and protein are interchangeable.

The polypeptides of the present invention are preferably provided in an isolated form, and may

be purified, *e.g.* to homogeneity. The term "isolated," when referring, *e.g.*, to a polypeptide or polynucleotide, means that the material is removed from its original environment (*e.g.*, the natural environment if it is naturally occurring), and isolated or separated from at least one other component with which it is naturally associated. For example, a naturally-occurring polypeptide present in its natural living host is not isolated, but the same polypeptide, separated from some or all of the coexisting materials in the natural system, is isolated. Such polypeptides could be part of a composition, and still be isolated as that composition is not part of its natural environment.

A fragment or variant of an TAFI polypeptide of the present invention preferably retains substantially at least one of the biological functions or activities of the polypeptide. Such a biological function or activity can be, *e.g.*, any of those described above, and includes having the ability to react with an antibody, *i.e.*, having an epitope-bearing peptide, particularly one which comprises one or more mutations of the invention. Fragments or variants of the polypeptides have sufficient similarity to those polypeptides so that at least one activity of the native polypeptides is retained. Fragments or variants of smaller polypeptides retain at least one activity (e.g., an enzymatic activity, or the ability to react with an antibody or antigen-binding fragment of the invention) of a comparable sequence found in the native polypeptide.

Polypeptide fragments of the invention may be of any size that is compatible with the invention. They may range in size from the smallest specific epitope (*e.g.*, about 6 contiguous amino acids from a particular sequence, more typically at least about 12 amino acids but can be up to 20 amino acids) to a nearly full-length gene product (*e.g.*, a single amino acid shorter than the full-length polypeptide (SEQ ID NO: 2)). For example, a polypeptide of the invention may comprise at lest about 10, 25, 50, 100, 200, 300, 400, 500, 600, 800, 1000, or 1200 amino acids.

Fragments of the polypeptides of the present invention may be employed, *e.g.*, for producing the corresponding full-length polypeptide by peptide synthesis, *e.g.*, as intermediates for producing the full-length polypeptides; for inducing the production of antibodies or antigen-binding fragments; as “query sequences” for the probing of public databases, or the like.

A variant of a polypeptide of the invention (*e.g.*, a variant of human TAFI which is already altered in any one of a number o sites thought to be involved in substrate binding) may be (i) one in which one or more of the amino acid residues are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code, or (ii) one in which one or more of the amino acid residues includes a substituent group, or (iii) one in which the polypeptide is fused with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol), or (iv) one in which additional amino acids are fused to the polypeptide, such as a leader or secretory sequence or a sequence which is employed for purification of the polypeptide, commonly for the purpose of creating a genetically engineered form of the protein that is susceptible to secretion from a cell, such as a transformed cell. The additional amino acids may be from a heterologous source, or may be endogenous to the natural gene.

Variant polypeptides belonging to type (i) above include, *e.g.*, muteins, analogs and derivatives. A variant polypeptide can differ in amino acid sequence by, *e.g.*, one or more additions, substitutions, deletions, insertions, inversions, fusions, and truncations or a combination of any of these. For example, conservative amino acid substitutions, which are well-known to those of skill in the art, generally do not lead to a change in protein function. Variant polypeptides belonging to type (ii) above include, *e.g.*, modified polypeptides. Known polypeptide modifications include, but are not limited to, glycosylation, acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent crosslinks, formation of cystine, formation of pyroglutamate, formylation, gamma carboxylation, glycosylation, GPI anchor formatin, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination.

Such modifications are well known to those of skill in the art and have been described in great detail in the scientific literature. Several particularly common modifications, glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues, hydroxylation and ADP-ribosylation, for instance, are described in many basic texts, such as *Proteins--Structure and Molecular Properties*, 2nd ed., T.E. Creighton, W.H. Freeman and Company, New York (1993). Many detailed reviews are available on this subject, such as by Wold, F., *Posttranslationail Covalent Modification of Proteins,* B.C. Johnson, Ed., Academic Press, New York 1-12 (1983); Seifter *et al.* (1990) *Meth. Enzymol.* 182:626-646 and Rattan *et al.* (1992) *Ann. N.Y. Acad. Sci.* 663:48-62.

Variant polypeptides belonging to type (iii) are well known in the art and include, *e.g.*, PEGylation or other chemical modifications.

Variants polypeptides belonging to type (iv) above include, *e.g.*, preproteins or proproteins, which can be activated by cleavage of the proprotein portion to produce an active mature polypeptide. Variants include a variety of hybrid, chimeric or fusion polypeptides. Typical examples of such variants are discussed elsewhere herein.

Many other types of variants are known to those of skill in the art. For example, as is well known, polypeptides are not always entirely linear. For instance, polypeptides may be branched as a result of ubiquitination, and they may be circular, with or without branching, generally as a result of post-translation events, including natural processing events and events brought about by human manipulation which do not occur naturally.

Variant polypeptides can be fully functional or can lack function in one or more activities, *e.g.*, in any of the functions or activities described above. Among the many types of useful variations are, *e.g.*, those which exhibit alteration of catalytic activity. For example, one embodiment involves a variation at the binding site that results in binding but not hydrolysis, or slower hydrolysis, of a substrate. A further useful variation at the same site can result in altered affinity for a substrate. Useful variations also include

changes that provide for affinity for another substrate. Another useful variation includes one that prevents activation by proteins such as trypsin. Another useful variation provides a fusion protein in which one or more domains or subregions are operationally fused to one or more domains or subregions from another carboxypeptidase. Another useful variation includes peptides with altered stability (*e.g.*, an altered half-life in a cell, tissue or host).

TAFI polypeptide mutants also include polypeptides that have varying degrees of sequence homology (identity) to a wild-type TAFI polypeptide or to a mutant TAFI polypeptide of the present invention. In one embodiment, the polypeptides are substantially homologous to a TAFI polypeptide of the present invention, or show substantial sequence homology (sequence identity) thereto. Thus, polypeptides, and fragments thereof, within the present invention may contain amino acid sequences which show at least about 65% -70% sequence homology (identity), preferably about 70-75%, 75%-80%, 80-85%, or 85-90% sequence homology (identity), and most preferably about 90-95% or 95-99% sequence homology (identity) to the polypeptides of the invention. The invention also encompasses polypeptides having a lower degree of sequence identity, but having sufficient similarity so as to perform one or more of the functions or activities exhibited by the carboxypeptidases of the invention.

In accordance with the present invention, the term "percent identity" or "percent identical," when referring to a sequence, means that a sequence is compared to a claimed or described sequence after alignment of the sequence to be compared (the "Compared Sequence") with the described or claimed sequence (the "Reference Sequence"). The Percent Identity is then determined according to the following formula:

Percent Identity = 100 [1-(C/R)]

wherein C is the number of differences between the Reference Sequence and the Compared Sequence over the length of alignment between the Reference Sequence and the Compared Sequence wherein (i) each base or amino acid in the Reference Sequence that does not have a corresponding aligned base or amino acid in the Compared Sequence and (ii) each gap in the Reference Sequence and (iii) each aligned base or amino acid in the Reference Sequence that is different from an aligned base or amino acid in the Compared Sequence, constitutes a difference; and R is the number of bases or amino acids in the Reference Sequence over the length of the alignment with the Compared Sequence with any gap created in the Reference Sequence also being counted as a base or amino acid.

In a preferred embodiment, the percent identity between two amino acid sequences is determined using the Needleman et al. (*J. Mol. Biol.*, 48:444-453 (1970)) algorithm. In yet another preferred embodiment, the percent identity between two nucleotide sequences is determined using the GAP program I the GCG software package (Devereux et al., *Nucleic Acids Res.*, 12 (1):387 (1984)). Additional algorithms for sequence analysis are known in the art and include, but are not limited to, ADVANCE and ADAM as described in Torellis et al., *Comput. Appl. Biosci.*, 10:3-5 (1994); FASTA described in Pearson et al., *Proc. Natl. Acad. Sci. USA,* 85:2444-8 (1988); and the algorithm of Myers and Miller, CABIOS (1989).

The description herein for percent identity or percent homology is intended to apply equally to

nucleotide or amino acid sequences

In accordance with the present invention, the term "substantially homologous," when referring to a protein sequence, means that the amino acid sequences are at least about 90-95% or 97-99% of the nucleotides or amino acids. A substantially homologous amino acid sequence of the invention can be encoded by a nucleic acid sequence hybridizing to the nucleic acid sequence, or portion thereof, of a sequence encoding a mutant polypeptide of theinvention, under conditions of high stringency.

Conditions of "high stringency," as used herein, means, for example, incubating a blot overnight (*e.g.*, at least 12 hours) with a long polynucleotide probe in a hybridization solution containing, *e.g.*, about 5X SSC, 0.5% SDS, 100 µg/ml denatured salmon sperm DNA and 50% formamide, at 42˚C. Blots can be washed at high stringency conditions that allow, *e.g.*, for less than 5% bp mismatch (*e.g.*, wash twice in 0.1X SSC and 0.1% SDS for 30 min at 65˚C), thereby selecting sequences having, *e.g.*, 95% or greater sequence identity.

Other non-limiting examples of high stringency conditions include a final wash at 65˚C in aqueous buffer containing 30 mM NaC1 and 0.5% SDS. Another example of high stringent conditions is hybridization in 7% SDS, 0.5 M $NaPO_4$, pH 7, 1 mM EDTA at 50˚C, *e.g.*, overnight, followed by one or more washes with a 1% SDS solution at 42˚C. Whereas high stringency washes can allow for less than 5% mismatch, reduced or low stringency conditions can permit up to 20% nucleotide mismatch. Hybridization at low stringency can be accomplished as above, but using lower formamide conditions, lower temperatures and/or lower salt concentrations, as well as longer periods of incubation time.

As used with respect to the polypeptides (and polynucleotides) of the present invention, the term fragment refers to a sequence that is a subset of a larger sequence (*i.e.*, a continuous or unbroken sequence of residues within a larger sequence). Fragments or larger peptides already present in the art are, of course, excluded.

Polynucleotides Encoding TAFI Polypeptide Mutants

The invention also includes polynucleotides, and fragments thereof, that encode, without interruption, the TAFI polypeptide mutants of the present invention. A polynucleotide that "codes without interruption" refers to a polynucleotide having a continuous open reading frame ("ORF") as compared to an ORF which is interrupted by introns or other noncoding sequences.

A polynucleotide of the present invention may be a recombinant polynucleotide, an isolated polynucleotide, or a synthetic or semi-synthetic polynucleotide, or combinations thereof. As used herein, the terms polynucleotide, oligonucleotide, oligomer and nucleic acid are interchangeable. Thus, reference to a "polynucleotide" can encompass fragments, such as oligonucleotides, of a full-length polynucleotide.

As used herein, the term "gene" means a segment of DNA involved in producing a polypeptide chain; it may include regions preceding and following the coding region (leader and trailer) as well as intervening sequences (introns) between individual coding segments (exons). cDNAs lack introns. The invention includes isolated genes (*e.g.*, genomic clones) which encode polypeptides of the invention.

Polynucleotides of the invention may be RNA, PNA, or DNA, *e.g.*, cDNA, genomic DNA, and synthetic or semi-synthetic DNA, or combinations thereof. The DNA may be triplex, double-stranded or single-stranded, and if single stranded, may be the coding strand or non-coding (anti-sense) strand. It can comprise hairpins or other secondary structures. The RNA includes oligomers (including those having sense or antisense strands), mRNAs, polyadenylated RNA, total RNA, single strand or double strand RNA, or the like. DNA/RNA duplexes are also encompassed by the invention.

The polynucleotides and fragments thereof of the present invention may be of any size that is compatible with the invention, *e.g.*, of any desired size that is effective to achieve a desired specificity when used as a probe. Polynucleotides may range in size, *e.g.*, from the smallest specific probe (*e.g.*, about 10-12 nucleotides) to greater than a full-length DNA, *e.g.*, in the case of a fusion polynucleotide or a polynucleotide that is part of a genomic sequence; fragments may be as large as, *e.g.*, one nucleotide shorter than a full-length DNA. A fragment of a polynucleotide according to the invention may be used, *e.g.*, as a hybridization probe, as discussed elsewhere herein.

Many types of variants of polynucleotides are encompassed by the invention including, *e.g.*, (i) one in which one or more of the nucleotides is substituted with another nucleotide, or which is otherwise mutated; or (ii) one in which one or more of the nucleotides is modified, *e.g.*, includes a substituent group; or (iii) one in which the polynucleotide is fused with another compound, such as a compound to increase the half-life of the polynucleotide; or (iv) one in which additional nucleotides are covalently bound to the polynucleotide, such a sequences encoding a leader or secretory sequence or a sequence which is employed for purification of the polypeptide. The additional nucleotides may be from a heterologous source, or may be endogenous to the natural gene.

Polynucleotide variants belonging to type (i) above include variants and mutants. Variant polynucleotides can comprise, *e.g.*, one or more additions, insertions, deletions, substitutions, transitions, transversions, inversions, chromosomal translocations, variants resulting from alternative splicing events, or the like, or any combinations thereof.

A polynucleotide of the invention may have a coding sequence which is a naturally or non-naturally –occurring allelic variant of a coding sequence encompassed by the sequence of wild-type TAFI. As is known in the art, an allelic variant is an alternate form of a polynucleotide sequence which may have a substitution, deletion or addition of one or more nucleotides, which in general does not substantially alter the function of the encoded polypeptide.

Other variant sequences, located in a coding sequence or in a regulatory sequence, may affect (e.g. enhance or decrease) the production of, or the function of activity of, a TAFI polypeptide mutant of the invention.

Polynucleotide variants belonging to type (ii) above include, *e.g.*, modifications such as the attachment of detectable markers or labels (avidin, biotin, radioactive or non-radioactive elements, fluorescent tags and dyes, energy transfer labels, energy-emitting labels, binding partners, etc.) or moieties which improve expression, uptake, cataloging, tagging, hybridization, detection, and/or stability. The polynucleotides can also be attached to solid supports, according to a desired method. See, *e.g.*,

U.S. Patent Nos. 5,470,967; 5,476,925; and 5,478,893.

Polynucleotide variants belonging to type (iii) above are well known in the art and include, *e.g.*, various lengths of polyA$^+$ tail, 5'cap structures, and nucleotide analogs, *e.g.*, inosine, thionucleotides, or the like.

Polynucleotide variants belonging to type (iv) above include, *e.g.*, a variety of chimeric, hybrid or fusion polynucleotides. For example, a polynucleotide of the invention can comprise a coding sequence and additional coding sequence (*e.g.*, sequences coding for leader, signal, secretory, targeting, enzymatic, fluorescent, antibiotic resistance, and other functional or diagnostic peptides); or a coding sequence and non-coding sequences (*e.g.*, untranslated sequences at either a 5' or 3' end, or dispersed in the coding sequence, *e.g.*, introns).

More specifically, the present invention includes polynucleotides wherein the coding sequence for the polypeptide is fused in the same reading frame to another polynucleotide sequence (*e.g.*, a leader sequence). A polypeptide having a leader sequence is a preprotein and may have the leader sequence cleaved by the host cell to form a mature form of the polypeptide. The polynucleotides may also encode for a proprotein which is the mature protein plus additional N-terminal amino acid residues. A mature protein having a prosequence is a proprotein and is generally an inactive form of the protein. Once the prosequence is cleaved an active protein remains.

Polynucleotides of the present invention may also have a coding sequence fused in frame to a marker sequence that allows for identification and/or purification of the polypeptide of the present invention. The marker sequence may be, *e.g.*, a hexa-histidine tag in the case of a bacterial host, or a hemagglutinin (HA) tag when a mammalian host is used.

Other types of polynucleotide variants will be evident to one of skill in the art. For example, the nucleotides of a polynucleotide can be joined via various known linkages, *e.g.*, ester, sulfamate, sulfamide, phosphorothioate, phosphoramidate, methylphosphonate, carbamate, etc., depending on the desired purpose, *e.g.*, resistance to nucleases, such as RNAse H, improved *in vivo* stability, etc. (See, *e.g.*, U.S. Patent No. 5,378,825). Any desired nucleotide or nucleotide analog can be incorporated, *e.g.*, 6-mercaptoguanine, 8-oxo-guanine, etc.

Also, polynucleotides of the invention may have a coding sequence derived from another genetic locus of an organism, providing it has a substantial homology to a mammalian wild-type TAFI polypeptide or to one from another organism (*e.g.*, an ortholog).

It is understood that variants exclude any sequences disclosed prior to the invention.

Polynucleotides according to the present invention can be labeled according to any desired method. The polynucleotide can be labeled using radioactive tracers such as, *e.g.*, ^{32}P, ^{35}S, ^{3}H, or ^{14}C by methods well known in the art. A non-radioactive labeling can also be used, combining a polynucleotide of the present invention with residues having immunological properties (antigens, haptens), a specific affinity for certain reagents (ligands), properties enabling detectable enzyme reactions to be completed (enzymes or coenzymes, enzyme substrates, or other substances involved in an enzymatic reaction), or characteristic physical properties, such as fluorescence or the emission or absorption of light at a desired

wavelength, etc.

The term "substantially homologous," when referring to polynucleotide sequences, means that the nucleotide sequences are at least about 90-95% or 95-99% or more identical.

Expression of TAFI Polypeptide Mutants

The present invention also relates to recombinant constructs that contain vectors plus polynucleotides of the present invention. Such constructs comprise a vector, such as a plasmid or viral vector, into which a polynucleotide sequence of the invention has been inserted, in a forward or reverse orientation.

Large numbers of suitable vectors are known to those of skill in the art, and many are commercially available. The following vectors are provided by way of example; Bacterial: pQE70, pQE60, pQE-9 (Qiagen), pBS, pD10, phagescript, psiX174, pBluescript SK, pBSKS, pNH8A, pNH16a, pNH18A, pNH46A (Stratagene); pTRC99a, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia); Eukaryotic: pWLNEO, pSV2CAT, pOG44, pXT1, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia). However, any other plasmid or vector may be used as long as it is replicable and viable in the host.

In a preferred embodiment, the vector is an expression vector, into which a polynucleotide sequence of the invention is inserted so as to be operably linked to an appropriate expression control (regulatory) sequence(s) (*e.g.*, promoters and/or enhancers) which directs mRNA synthesis. Appropriate expression control sequences, *e.g.*, regulatable promoter or regulatory sequences known to control expression of genes in prokaryotic or eukaryotic cells or their viruses, can be selected for expression in prokaryotes (*e.g.*, bacteria), yeast, plants, mammalian cells or other cells. Preferred expression control sequences are derived from highly-expressed genes, *e.g.*, from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), α-factor, acid phosphatase, or heat shock proteins, among others. Such expression control sequences can be selected from any desired gene, *e.g* using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors for such selection are pKK232-8 and pCM7.

Particular named bacterial promoters, which can be used include lacI, lacZ, T3, T7, gpt, lambda P_R, P_L and trp. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, adenovirus promoters, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is within the level of ordinary skill in the art.

Transcription of the DNA encoding the polypeptides of the present invention by higher eukaryotes can be increased by inserting an enhancer sequence into the expression vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp that act on a promoter to increase its transcription. Representative examples include the SV40 enhancer on the late side of the replication origin bp 100 to 270, a cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

Generally, recombinant expression vectors also include origins of replication. An expression vector may contain a ribosome binding site for translation initiation, a transcription termination sequence,

a polyadenylation site, splice donor and acceptor sites, and/or 5' flanking or non-transcribed sequences. DNA sequences derived from the SV40 splice and polyadenylation sites may be used to provide required nontranscribed genetic elements. The vector may also include appropriate sequences for amplifying expression. In addition, expression vectors preferably contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells such as dihydrofolate reductase or neomycin resistance for eukaryotic cell culture, or such as tetracycline or ampicillin resistance in *E. coli*.

Large numbers of suitable expression vectors are known to those of skill in the art, and many are commercially available. Suitable vectors include chromosomal, nonchromosomal and synthetic DNA sequences, *e.g.*, derivatives of SV40; bacterial plasmids; phage DNA; baculovirus; yeast plasmids; vectors derived from combinations of plasmids and phage DNA, viral DNA such as vaccinia, adenovirus, adeno-associated virus, TMV, fowl pox virus, and pseudorabies. However, any other vector may be used as long as it is replicable and viable in a host. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described, *e.g.*, by Sambrook, et al., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, N.Y., (1989), Wu et al., *Methods in Gene Biotechnology* (CRC Press, New York, NY, 1997), *Recombinant Gene Expression Protocols*, in *Methods in Molecular Biology*, Vol. 62, (Tuan, ed., Humana Press, Totowa, NJ, 1997), and *Current Protocols in Molecular Biology,* (Ausabel et al., Eds.,), John Wiley & Sons, NY (1994-1999).

In a preferred embodiment, a Baculovirus-based expression system is used. Baculoviruses represent a large family of DNA viruses that infect mostly insects. The prototype is the nuclear polyhedrosis virus (AcMNPV) from *Autographa californica,* which infects a number of lepidopteran species. One advantage of the baculovirus system is that recombinant baculoviruses can be produced *in vivo*. Following co-transfection with transfer plasmid, most progeny tend to be wild type and a good deal of the subsequent processing involves screening. To help identify plaques, special systems are available that utilize deletion mutants. By way of non-limiting example, a recombinant AcMNPV derivative (called BacPAK6) has been reported in the literature that includes target sites for the restriction nuclease Bsu36I upstream of the polyhedrin gene (and within ORF 1629) that encodes a capsid gene (essential for virus viability). Bsf36I does not cut elsewhere in the genome and digestion of the BacPAK6 deletes a portion of the ORF1629, thereby rendering the virus non-viable. Thus, with a protocol involving a system like Bsu36I-cut BacPAK6 DNA most of the progeny are non-viable so that the only progeny obtained after co-transfection of transfer plasmid and digested BacPAK6 is the recombinant because the transfer plasmid, containing the exogenous DNA, is inserted at the Bsu36I site thereby rendering the recombinants resistant to the enzyme. See Kitts and Possee, A method for producing baculovirus expression vectors at high frequency, *BioTechniques*, 14, 810-817 (1993). For general procedures, see King and Possee, *The Baculovirus Expression System: A Laboratory Guide*, Chapman and Hall, New York (1992) and *Recombinant Gene Expression Protocols*, in *Methods in Molecular Biology*, Vol. 62, (Tuan, ed., Humana Press, Totowa, NJ, 1997), at Chapter 19, pp. 235-246.

Appropriate DNA sequences may be inserted into a vector by any of a variety of procedures. In general, the DNA sequence is inserted into an appropriate restriction endonuclease site(s) by procedures

known in the art. Such procedures and others are deemed to be within the scope of those skilled in the art. Conventional procedures for this and other molecular biology techniques discussed herein are found in many readily available sources, *e.g.*, Sambrook, *et al.*, *Molecular Cloning: A Laboratory Manual,* Second Edition, Cold Spring Harbor, N.Y., (1989). If desired, a heterologous structural sequence is assembled in an expression vector in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium.

The present invention also relates to host cells which are transformed/transfected/transduced with constructs such as those described above, and to progeny of said cells, especially where such cells result in a stable cell line that can be used for assays of carboxypeptidase activity, *e.g.*, in order to identify agents which modulate carboxypeptidase activity, and/or for production (*e.g.*, preparative production) of the polypeptides of the invention.

As representative examples of appropriate hosts, there may be mentioned: bacterial cells, such as *E. coli*, *Streptomyces*, *Salmonella typhimurium*; fungal cells, such as yeast; insect cells such as *Drosophila S2* and *Spodoptera Sf9* (and other insect expression systems); animal cells, including mammalian cells such as CHO, QM7, COS (*e.g.*, the COS-7 lines of monkey kidney fibroblasts described by Gluzman, *Cell*, 23:175 (1981)), C127, 3T3, CHO, HeLa, BHK or Bowes melanoma cell lines; plant cells, etc. The selection of an appropriate host is deemed to be within the knowledge of those skilled in the art based on the teachings herein. Cell lines used for testing putative modulatory agents are commonly mammalian cells.

In a preferred embodiment, the host cells are insect cells of *Spodoptera* species, most especially SF9 cells, from *Spodoptera frugiperda*. Polypeptides (*e.g.*, full length polypeptides) of the present invention are readily obtainable from insect cells using a baculovirus expression vector. Such expression is readily characterized using methods well known in the art.

Introduction of a construct into a host cell can be effected by, *e.g.*, calcium phosphate transfection, DEAE-Dextran mediated transfection, lipofection a gene gun, or electroporation (Davis, L., Dibner, M., Battey, I., *Basic Methods in Molecular Biology*, (1986)).

Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter can be induced by appropriate means (*e.g.*, temperature shift or chemical induction) if desired, and cells cultured for an additional period. The engineered host cells are cultured in conventional nutrient media modified as appropriate for activating promoters (if desired), selecting transformants or amplifying the genes of the present invention. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification. Alternatively, when a heterologous polypeptide is secreted from the host cell into the culture fluid, supernatants of the culture fluid can be used as a source of the protein. Microbial cells employed in expression of proteins can be disrupted by any convenient

method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents, such methods being well known to those skilled in the art.

The polypeptide can be recovered and purified from recombinant cell cultures by conventional methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography, or the like. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. High performance liquid chromatography (HPLC) can be employed for final purification steps.

In addition to the methods described above for producing polypeptides recombinantly from a prokaryotic or eukaryotic host, polypeptides of the invention can be prepared from natural sources, or can be prepared by chemical synthetic procedures (*e.g.*, synthetic or semi-synthetic), *e.g.*, with conventional peptide synthesizers. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention.

Detection and/or quantitation of protein levels can be accomplished by any of a variety of conventional methods, e.g., methods based on antibodies or antigen-specific fragments of the invention. Immunological assays include, e.g., ELISA, RIA and FACS assays. A two-site, monoclonal-based immunoassay, utilizing antibodies reactive to two non-interfering epitopes on a polypeptide of the invention are preferred, but a competitive binding assay may be employed. These and other assays are described, e.g., in Hampton et al., *Serological Methods, a Laboratory Manual*, APS Press, St. Paul, Minn. (1990).

Proteins of the invention can also be expressed in, and isolated and/or purified from, transgenic animals or plants. Procedures to make and use such transgenic organisms are standard in the art. Some such procedures are described elsewhere herein.

Screening for substrates and modulatory agents for TAFI and related carboxypeptidases

TAFI polypeptide mutants of the invention are useful for screening agents, *in vitro* or *in vivo* (*e.g.*, in cell-based assays or in animal models), to identify agents that are novel substrates /inhibitors/modulators of TAFI, CPN or other related carboxypeptidases.

In screening for potential substrates/inhibitors/modulators of TAFI polypeptide mutants or related carboxypeptidases, enzymatic activities normally associated with the wild-type TAFI (SEQ ID NO: 2) can be employed. Assays can be performed *in vitro*, *ex vivo* or *in vivo*. In a preferred embodiment, such assays can be performed using isolated TAFI polypeptide mutants of the invention. Standard assays are available to measure (*e.g.*, quantitate) levels of carboxypeptidase activity (Hendriks et al., *Clinica. Chimica Acta*, 157:103-8 (1986); Zhao et al., *Thromb. Haemost.,* 80:949-55 (1998)), as demonstrated in Examples III and IV. This activity can be compared to the activity of a reference molecule, such as wild-type TAFI or to another carboxypeptidase of interest. A decrease in carboxypeptidase activity indicates inhibitory activity by the agent being tested, while an increase in TAFI or carboxypeptidase activity indicates an activating effect by the agent being tested.

In another embodiment, stable cell lines, such as CHO or 293 cells that express a mutant polypeptide of the invention, are treated with a putative modulatory agent, and the carboxypeptidase activity is measured.

Other standard methods can be used to measure the binding affinity of putative inhibitors, stimulators or substrates of the polypeptides of the invention or other related carboxypeptidases, or to measure the ability of a putative inhibitor or enhancer to stimulate or inhibit interactions between the TAFI polypeptide mutants of the invention or other related carboxypeptidases and a target molecule. An example of an assay for an antagonist combines a polypeptide of the invention and a potential antagonist (*i.e.*, an inhibitor) under appropriate conditions for a competitive inhibition assay.

Any of the assays described herein can, of course, be adapted to any of a variety of high throughput methodologies, as can the generation, identification and characterization of putative inhibitory or stimulatory agents. Agents identified on the basis of their ability to modulate expression or the TAFI polypeptide mutants of the invention, may also be used for modulating TAFI or other carboxypeptidases, and/or for diagnosing or treating disease conditions related thereto.

The invention also provides methods of screening agents, *in vitro* or *in vivo* (*e.g.,* in cell-based assays or in animal models), to identify those agents that modulate (*e.g.*, enhance, stimulate, restore, inhibit, block, stabilize, destabilize, increase, facilitate, up-regulate, activate, amplify, augment, induce, decrease, down-regulate, diminish, lessen, reduce, etc.) synthesis and/or activity of the TAFI polypeptide mutants of the invention. Agents that inhibit such synthesis and/or activity (antagonists) may, *e.g.*, result in reduced thrombotic tendency. Agents that enhance such synthesis and/or activity (agonists) may, *e.g.*, result in improved wound healing.

Agents which inhibit the TAFI polypeptide mutants of the invention or other related carboxypeptidase activity (sometimes referred to herein as "TAFI or carboxypeptidase inhibitors") can be used to treat, prevent, and/or ameliorate the symptoms of conditions associated with an overexpression or increased activity of TAFI or carboxypeptidase; and agents which enhance such activity can be used to treat, prevent, and/or ameliorate the symptoms of conditions associated with an underexpression or decreased activity of TAFI or carboxypeptidase. Inhibitors of TAFI or related carboxypeptidases can be used, *e.g.*, to treat any of the conditions described elsewhere herein which are associated with an overproduction of or increased activity of TAFI or carboxypeptidase. Stimulators of TAFI or carboxypeptidase can be used, *e.g.*, to treat any of the conditions described elsewhere herein which are associated with an underproduction of or decreased activity of TAFI, the polypeptides of the invention or other related carboxypeptidases.

Potential modulators, *e.g.*, inhibitors or activators, of the invention, include, *e.g.*, small chemical compounds (*e.g.*, inorganic or organic molecules), polypeptides, peptides or peptide analogs, polynucleotides, antibodies that bind specifically to the polypeptides of the invention, or the like. Typical polypeptide agents include, *e.g.*, mutant polypeptides of the invention or fragments thereof which exhibit impaired enzymatic activity but which have a higher affinity for a target than the polypeptides of the invention; such polypeptides can out compete the polypeptides of the invention and, thus, inhibit its

activity. Other inhibitory or stimulatory substances may enter cells and bind directly to the DNA neighboring the sequences coding for the polypeptides of the invention, thereby decreasing or increasing their expression.

One class of modulators includes small molecules that bind to and occupy the catalytic site of the polypeptide, thereby making the catalytic site inaccessible to a substrate such that normal biological activity is prevented. Catalytic sites can be determined by conventional, art-recognized methods, *e.g.*, comparison to catalytic sites found in related carboxypeptidases. For example, carboxypeptidases often include the catalytic signature site comprising residues Glu271 and Arg125.

Antibodies, antigen-binding fragments or other specific binding partners

The polypeptides, their fragments or variants thereof, or cells expressing them can also be used as immunogens to produce specific antibodies, or antigen-binding fragments, thereto. By a "specific" antibody or antigen-binding fragment is meant one, which binds selectively (preferentially) to a polypeptide of the invention, or to a fragment or variant thereof. An antibody "specific" for a polypeptide means that the antibody recognizes a defined sequence of amino acids within or including the polypeptide.

Antibodies of the invention can be, for example, polyclonal or monoclonal antibodies. The present invention also includes chimeric, recombinant, single chain, and partially or fully humanized antibodies, as well as Fab fragments, or the product of a Fab expression library, and fragments thereof. The antibodies can be IgM, IgG, subtypes, IgG2A, IgG1, etc. Various procedures known in the art may be used for the production of such antibodies and fragments. Antibodies generated against the polypeptides corresponding to a sequence of the present invention can be obtained, *e.g.*, by direct injection of the polypeptides into an animal or by administering the polypeptides to an animal, *e.g.*, goat, rabbit, mouse, chicken, etc., preferably a non-human. The antibody so obtained will then bind the polypeptide itself. In this manner, even a sequence encoding only a fragment of the polypeptides can be used to generate antibodies binding the whole native polypeptides. Such antibodies can then be used to isolate the polypeptide from tissue expressing that polypeptide. Antibodies can also be generated by administering naked DNA. See, *e.g.*, U.S. Patent Nos. 5,703,055; 5,589,466; and 5,580,859.

For preparation of monoclonal antibodies, any technique, which provides antibodies produced by continuous cell line cultures can be used. Examples include, *e.g.*, the hybridoma technique (Kohler and Milstein, *Nature*, 256:495-7 (1975)), the trioma technique, the human B-cell hybridoma technique (Kozbor et al., *Immunology Today*, 4:72 (1983)), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole et al., *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96 (1985)).

Techniques described for the production of single chain antibodies (*e.g.*, U.S. Patent No. 4,946,778) can be adapted to produce single chain antibodies to immunogenic polypeptide products of this invention. Also, transgenic animals may be used to express partially or fully humanized antibodies to immunogenic polypeptide products of this invention.

The invention also relates to other specific binding partners, which include, *e.g.*, aptamers and PNA.

In accordance with the present invention, an antibody or antigen-binding fragment can be present in a kit, where the kit includes, *e.g.*, one or more antibodies or antigen-binding fragments, a desired buffer, detection compositions, proteins (*e.g.*, wild type) to be used as controls, etc.

Transgenic and knock-in animals

The invention disclosed herein also relates to a non-human transgenic animal comprising within its genome one or more copies of the polynucleotides encoding the novel polypeptides of the invention. The transgenic animals of the invention may contain within their genome multiple copies of the polynucleotides encoding the polypeptides of the invention, or one copy of a gene encoding such polypeptide but wherein said gene is linked to a promoter (*e.g.*, a regulatable promoter) that will direct expression (preferably overexpression) of said polypeptide within some, or all, of the cells of said transgenic animal. In a preferred embodiment, expression of a polypeptide of the invention occurs preferentially in liver tissue. Other tissues where the polypeptides of the invention can be expressed include, but are not limited to, brain, lung, and kidney. A variety of non-human transgenic organisms are encompassed by the invention, including *e.g.*, drosophila, C. elegans, zebrafish and yeast. The transgenic animal of the invention is preferably a mammal, *e.g.*, a cow, goat, sheep, rabbit, non-human primate, or rat, most preferably a mouse.

Methods of producing transgenic animals are well within the skill of those in the art, and include, *e.g.*, homologous recombination, mutagenesis (*e.g.*, ENU, Rathkolb et al., *Exp. Physiol.*, 85(6):635-44, (2000)), and the tetracycline-regulated gene expression system (*e.g.*, U.S. Patent No. 6,242,667), and will not be described in detail herein. See *e.g.*, Wu et al., *Methods in Gene Biotechnology*, CRC, pp.339-366 (1997); Jacenko, O., Strategies in Generating Transgenic Animals, *Recombinant Gene Expression Protocols*, Vol. 62 of *Methods in Molecular Biology*, Humana Press, pp. 399-424 (1997).

Transgenic organisms are useful, *e.g.*, for providing a source of a polynucleotide or polypeptide of the invention, or for identifying and/or characterizing agents that modulate expression and/or activity of such a polynucleotide or polypeptide. Transgenic animals are also useful as models for disease conditions related to, *e.g.*, overexpression of a polynucleotide or polypeptide of the invention.

The present invention also relates to a transgenic non-human animal whose genome comprises one or more genes coding for the polypeptides disclosed herein in place of the mammalian gene otherwise coding for said non-human isoform. Most preferably said animal is a mouse.

The present invention also relates to a non-human knock-in animal whose genome lacks or fails to express a functional endogenous polypeptide (*e.g.*, an endogenous TAFI or CPN polypeptide), but instead expresses a polypeptide of the invention or functional analog thereof (*i.e.*, the endogenous gene is functionally disrupted and a polynucleotide of the invention is expressed in its place).

Functional disruption of the endogenous gene can be accomplished in any effective way, including, *e.g.*, introduction of a stop codon into any part of the coding sequence such that the resulting polypeptide is biologically inactive (*e.g.*, because it lacks a catalytic domain, a ligand binding domain, etc.), introduction of a mutation into a promoter or other regulatory sequence that is effective to turn it off, or reduce transcription of the gene, insertion of an exogenous sequence into the gene which inactivates it (*e.g.*, which disrupts the production of a biologically-active polypeptide or which disrupts the promoter or other transcriptional machinery), deletion of sequences from the endogenous gene, etc. Examples of transgenic animals having functionally disrupted genes are well known, *e.g.*, as described in U.S. Patent Nos. 6,239,326, 6,225,525, 6,207,878, 6,194,633, 6,187,992, 6,180,849, 6,177,610, 6,100,445, 6,087,555, 6,080,910, 6,069,297, 6,060,642, 6,028,244, 6,013,858, 5,981,830, 5,866,760, 5,859,314, 5,850,004, 5,817,912, 5,789,654, 5,777,195, and 5,569,824.

For creating functional disrupted genes, and other gene mutations, homologous recombination technology is of special interest since it allows specific regions of the genome to be targeted. Using homologous recombination methods, genes can be specifically inactivated, specific mutations can be introduced, and exogenous sequences can be introduced at specific sites. These methods are well known in the art, *e.g.*, as described in the patents above. See also, Robertson, *Biol. Reproduc.*, 44(2):238-45, (1991). Generally, the genetic engineering is performed in an embryonic stem (ES) cell, or other pluripotent cell line (*e.g.*, adult stem cells, EG cells), and that genetically-modified cell (or nucleus) is used to create a whole organism. Nuclear transfer can be used in combination with homologous recombination technologies.

For example, a locus of TAFI or a related carboxypeptidase can be disrupted in mouse ES cells using a positive-negative selection method (*e.g.*, Mansour et al., *Nature*, 336:348-52 (1988)). In this method, a targeting vector can be constructed which comprises a part of the gene to be targeted. A selectable marker, such as neomycin resistance genes, can be inserted into a TAFI exon present in the targeting vector, disrupting it. When the vector recombines with the ES cell genome, it disrupts the function of the gene. The presence in the cell of the vector can be determined by expression of neomycin resistance. See, *e.g.*, U.S. Patent No. 6,239,326. Cells having at least one functionally disrupted gene can be used to make chimeric and germline animals, *e.g.*, animals having somatic and/or germ cells comprising the engineered gene. Homozygous knock-in animals can be obtained from breeding heterozygous knock-in animals.

A knock-in animal, or animal cell, lacking one or more functional genes encoding TAFI (or a related carboxypeptidase) but expressing one or more of the polypeptides of the invention can be useful in a variety of applications, including drug screening assays (*e.g.,* for carboxypeptidases other than TAFI; by making a cell deficient in TAFI, the contribution of other carboxypeptidases can be specifically examined), as a source of tissues deficient in TAFI activity, as the starting material for generating an animal in which the endogenous TAFI is replaced with human TAFI or the polynucleotides of the invention, and any of the utilities mentioned in any issued U.S. Patent on transgenic animals.

In addition to the methods mentioned above, transgenic or knock-in animals can be prepared according to known methods in the art, including, *e.g.*, by pronuclear injection of recombinant genes into pronuclei of 1-cell embryos, incorporating an artificial yeast chromosome into embryonic stem cells, gene targeting methods, embryonic stem cell methodology, cloning methods, nuclear transfer methods.

A polynucleotide according to the present invention can be introduced into any non-human animal, including a non-human mammal, mouse, sheep, cattle, rat, or primate. Transgenic animals can be produced by the methods described in U.S. Patent No. 5,994,618, and utilized for any of the utilities described therein.

Therapeutics

The methods of the present invention are also directed to facilitating the development of potentially useful therapeutic agents that may be effective in combating TAFI mediated or related disease conditions, and to methods of effecting such treatments.

Any agent which modulates the expression and/or activity of the TAFI polypeptide mutants or the polynucleotides encoding them identified by an art recognized assay, such as those herein, can be used therapeutically. Some such agents are discussed elsewhere herein.

Agents which affect expression and/or activities of polypeptides of the invention can be administered to patients in need thereof by conventional procedures, in order to prevent or treat conditions as disclosed elsewhere herein and/or to ameliorate symptoms of those conditions. Such agents can be formulated into pharmaceutical compositions comprising pharmaceutically acceptable excipients, carriers, etc., using conventional methodologies. Formulations and excipients, which enhance transfer (promote penetration) of an agent across the blood-brain barrier are also well known in the art.

In addition to agents, which can moderate the expression or activity of a carboxypeptidase of the invention, treatment methods according to the invention also encompass the administration of a carboxypeptidase of the invention or variant or fragment thereof to a patient in need of such therapy. For example, such a polypeptide or fragment can compensate for reduced or aberrant expression or activity of the protein, and/or, by virtue of, *e.g.*, higher affinity for a target or substrate, can provide effective competition for it. In another embodiment, conventional methods of immunotherapy can be used.

Polynucleotides of the invention can also be used in methods of gene therapy, *e.g.*, utilized in gene delivery vehicles. The gene delivery vehicle may be of viral or non-viral origin. See generally, Jolly, *Cancer Gene Therapy,* 1:51-64 (1994); Kimura, *Human Gene Therapy,* 5:845-52 (1994); Connelly, *Human Gene Therapy,* 1:185-93 (1995); and Kaplitt, *Nature Genetics,* 6:148-153 (1994). Gene therapy vehicles for delivery of constructs including a coding sequence of a therapeutic of the invention can be administered either locally or systemically. These constructs can utilize viral or non-viral vector approaches. Expression of such coding sequences can be induced using endogenous mammalian or heterologous promoters. Expression of the coding sequence can be either constitutive

or regulated.

The present invention can employ recombinant retroviruses, which are constructed to carry or express a selected nucleic acid molecule of interest. Retrovirus vectors that can be employed are well known in the art. Preferred recombinant retroviruses include those described in WO 91/02805. Packaging cell lines suitable for use with the above-described retroviral vector constructs may be readily prepared (see PCT publications WO 95/30763 and WO 92/05266), and used to create producer cell lines (also termed vector cell lines) for the production of recombinant vector particles.

The present invention also employs alphavirus-based vectors that can function as gene delivery vehicles. Such vectors can be constructed from a wide variety of alphaviruses, including, for example, Sindbis virus vectors, Semliki forest virus, Ross River virus and Venezuelan equine encephalitis virus. Gene delivery vehicles of the present invention can also employ parvovirus such as adeno-associated virus (AAV) vectors. Exemplary adenoviral gene therapy vectors employable in this invention include those described in WO 94/12649, WO 93/03769; WO 93/19191; WO 94/28938; WO 95/11984 and WO 95/00655. Administration of DNA linked to killed adenovirus as described in Curiel, *Hum. Gene Ther.*, 3:147-54 (1992), may be employed.

Naked DNA may also be employed. Exemplary naked DNA introduction methods are described in WO 90/11092 and U.S. Patent No. 5,580,859. Liposomes that can act as gene delivery vehicles are described in U.S. Patent No. 5,422,120, PCT Patent Publication Nos. WO 95/13796, WO 94/23697 and WO 91/14445, and EP No. 0 524 968.

Other gene delivery vehicles and methods well known in the art may be employed. For example, non-viral delivery suitable for use includes mechanical delivery systems such as the approach described in Woffendin et al., *Proc. Natl. Acad. Sci.* USA, 91(24):11581-5 (1994). Moreover, the coding sequence and the product of expression of such can be delivered through deposition of photopolymerized hydrogel materials. Other conventional methods for gene delivery that can be used for delivery of the coding sequence include, for example, use of hand-held gene transfer particle gun, as described in U.S. Patent No. 5,149,655; use of ionizing radiation for activating transferred gene, as described in U.S. Patent No. 5,206,152 and PCT Patent Publication No. WO 92/11033.

Computer-based applications

The nucleotide or amino acid sequences of the invention are also provided in a variety of media to facilitate use thereof. As used herein, "provided" refers to a manufacture, other than an isolated nucleic acid or amino acid molecule, which contains a nucleotide or amino acid sequence of the present invention. Such a manufacture provides the nucleotide or amino acid sequences, or a subset thereof (*e.g.*, a subset of open reading frames (ORFs)) in a form which allows a skilled artisan to examine the manufacture using means not directly applicable to examining the nucleotide or amino acid sequences, or a subset thereof, as they exist in nature or in purified form.

In one application of this embodiment, a nucleotide or amino acid sequence of the present

invention can be recorded on computer readable media. As used herein, "computer readable media" refers to any medium that can be read and accessed directly by a computer. The skilled artisan will readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising computer readable medium having recorded thereon a nucleotide or amino acid sequence of the present invention.

As used herein, "recorded" refers to a process for storing information on computer readable medium. The skilled artisan can readily adopt any of the presently known methods for recording information on computer readable medium to generate manufactures comprising the nucleotide or amino acid sequence information of the present invention.

A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide or amino acid sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen to access the stored information. The skilled artisan can readily adapt any number of dataprocessor structuring formats (*e.g.*, text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention.

By providing the nucleotide or amino acid sequences of the invention in computer readable form, the skilled artisan can routinely access the sequence information for a variety of purposes. For example, one skilled in the art can use the nucleotide or amino acid sequences of the invention in computer readable form to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of the sequences of the invention, which match a particular target sequence or target motif. As used herein, a "target sequence" can be any DNA or amino acid sequence of six or more nucleotides or two or more amino acids. The most preferred sequence length of a target sequence is from about 10 to 100 amino acids or from about 30 to 300 nucleotide residues. However, it is well recognized that commercially important fragments, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

As used herein, "a target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequence(s) are chosen on a three-dimensional configuration, which is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzyme active sites, substrate binding sites and signal sequences. Nucleic acid target motifs include, but are not limited to, promoter sequences, hairpin structures and inducible expression elements (protein binding sequences).

The skilled artisan can use a particular three-dimensional configuration (*e.g.*, enzyme active sites or substrate binding sites), which is formed upon the folding of the target motif, for rational drug design or to design substrates or other molecules that may alter the function or activity of TAFI, CPN, or other related carboxypeptidases. Methods for rational drug design or the design of small molecules are well known in the art.

Computer software is publicly available which allows a skilled artisan to access sequence

information provided in a computer readable medium for analysis and comparison to other sequences. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are and can be used in the computer-based systems of the present invention. Examples of such software include, but are not limited to, MacPattern (EMBL), BLASTN and BLASTX (NCBIA).

For example, software, which implements the BLAST (Altschul et al., *J. Mol. Biol.*, 215:403-10 (1990)) and BLAZE (Brutlag et al., *Comp. Chem.*, 17:203-7 (1993)) search algorithms on a Sybase system can be used to identify open reading frames (ORFs) of the sequences of the invention which contain homology to ORFs or proteins from other libraries. In addition, the nucleotide or amino acid sequences of the invention as provided herein can be used to search databases of other nucleotide or amino acid sequences.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

In the foregoing and in the following examples, all temperatures are set forth uncorrected in degrees Celsius; and, unless otherwise indicated, all parts and percentages are by weight.

EXAMPLE I: Construction of TAFI polypeptide mutants

All polypeptide mutants of the invention were constructed based on the previously described wild type sequence (Eaton et al., *J. Biol. Chem.*, 266:21833-8, (1991)), except with Ala at position 147 and Thr at position 325 of the TAFI zymogen (Zhao et al., *Thromb. Haemost.*, 80:949-55 (1998)). Amino acid changes were introduced into pCEP 4/TAFI, a plasmid coding for wild type TAFI cDNA using a QuickChange Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA) according to the manufacturer's instructions. A pair of primers was used to construct each of four mutant polypeptides, each containing one altered amino acid. One polypeptide contains an Asp to Gln mutation at position 256 (TAFI D256Q) (numbering starts with the first alanine of activated TAFI (SEQ ID NO: 3)), one with Ser to Asp mutation at position 207 (S207D), one with Leu to Trp mutation at position 248 (L248W), and one with Asp to Ala mutation at position 257 (D257A). The primers used for

D256Q were 5'-ATAGATCCAATCCTGCCCACCTCCAGG-3' (SEQ ID NO: 4) and
5'-CCTGGAGGTGGGCAGGATTGGATCTAT-3' (SEQ ID NO: 5);

for S207D were 5'-GGGACGATTGGATCTATCTATTGGGCATCAAATATTC-3' (SEQ ID NO: 6) and
5'-GAATATTTGATGCC CAATAGATAGATCCAATCGTCCC-3' (SEQ ID NO: 7);

for L248W were 5'-GGCTCAGAAACCTGGTACCTAGCTCCTG-3' (SEQ ID NO: 8) and
5'-CAGGAGCTAGGTACCAGGTTT CTGAGCC-3' (SEQ ID NO: 9);

and for D257A were 5'-TGGAGGTGGGGATGCTTGGATCTATGA-3' (SEQ ID NO: 10) and 5'-TCATAGATCCAAGCATCCCCACCTCCA-3' (SEQ ID NO: 11).

A polypeptide mutant containing a double mutant, D256Q/S207D, was constructed with the same pair of primers (SEQ ID NOS: 6 and 7) used to change Ser 207 to Asp, using D256Q DNA as a template. The amino acid changes were confirmed by DNA sequence analysis.

Table 1: A summary of the mutations made in human TAFI and corresponding residues found in CPN and CPA. The residue number is based on SEQ ID NO: 3 for activated TAFI (TAFIa).

Residue	*TAFI wt*	*TAFI mutants*	*CPN*	*CPA*
207	*Ser*	*Asp*	*Asp*	*Gly*
248	*Leu*	*Trp*	*Trp*	*Ile*
256	*Asp*	*Gln*	*Gln*	*Ile*
257	*Asp*	*Ala*	*Asp*	*Asp*

EXAMPLE II: Expression and Purification of wild type and mutant TAFI polypeptides

Expression plasmids coding for wild type TAFI and the TAFI polypeptide mutants were transfected into 293 cells using Lipofectin (GibcoBRL). Cell lines with stable expression of TAFI and the polypeptides of the invention were selected by culturing cells in the presence of 150 μg/ml Hygromycin B. F or production of recombinant wild type and mutant TAFI polypeptides, stably expressing lines were cultured in Opti-MEM (GibcoBRL). Conditioned media was harvested either at 48 h or 72 h intervals. Harvested media was filtered through a 0.22-μm filter and stored at -20°C. Expression of wild-type and mutant TAFI polypeptides in the conditioned medium was confirmed by separation of the proteins on SDS/PAGE followed by Western blotting with polyclonal sheep-antiTAFI-antibody (Enzyme Research Lab).

To isolate wild type TAFI and the TAFI mutant polypeptides, typically 1 liter of conditioned medium was diluted 20-fold with H_2O, and adjusted to pH 6.8 with 0.1 M HCl. The sample was applied to a 25 ml SP-Sepharose column (Amersham Biosciences) equilibrated with 20 mM phosphate buffer, pH 6.8, and washed extensively with the same buffer. Bound proteins were eluted from the column with a salt gradient of 0 to 0.5 M NaCl in 20 mM phosphate buffer, pH 6.8. Fractions were analyzed by Western blotting as above. Fractions containing the polypeptides of interest were pooled, and concentrated using centriprep 10 (Amicon). The purity of the protein was determined by silver-stained

SDS/PAGE gel electrophoresis. The quantity of the polypeptides was determined both by using the TAFI ELISA kit (Enzyme Research Lab) as well as western blotting with polyclonal sheep anti-TAFI antibody.

Recombinant polypeptides were identified using silver staining and Western blotting with polyclonal sheep-antiTAFI-antibody after separation on SDS/PAGE (Fig. 2). The molecular weight of both wild type TAFI and the polypeptides of the invention is approximately 58 kDa, which is similar to TAFI purified from plasma.

EXAMPLE III: Activation of TAFI and Carboxypeptidase B Activity Assay

Recombinant wild type TAFI or mutant TAFI polypeptides (96 ng) were activated with 15 nM of thrombin (American Diagnostica), 50 nM thrombomodulin (Glaser et al., *J. Clin. Invest.*, 90: 2565-73, (1992)), and 5 mM of $CaCl_2$ in 20 mM HEPES, 0.15 M NaCl, pH 7.4, at room temperature for 5 min to 20 min. Activation was stopped by addition of 1 μM PPACK (Calbiochem), and was confirmed by analyzing an aliquot of the reaction by SDS/PAGE electrophoresis of denatured samples, followed by Western blotting using polyclonal sheep-antiTAFI-Ig antibody. A non-activated control was performed by preincubating PPACK with thrombin for 10 min before recombinant wild type TAFI or the polypeptides of the invention and thrombomodulin were added.

Carboxypeptidase B-like activity of the activated polypeptides was measured in a chromogenic assay using TAFI Developer (American Diagnostica Inc.), a specific TAFI substrate. In a 96-well microtiter plate, 25 μl of each activated polypeptide (1μg/ml) from the reactions above were mixed with 125 μl of 20 mM HEPES, 0.15 M NaCl, pH 7.4, and 50 μl of TAFI Developer for 30 min with shaking at 37°C. The reactions were stopped by the addition of 50 μl of 2 M sulfuric acid, and absorbance of the solutions was read in a spectrophotometer at 490 nm. Twenty five μl of carboxypeptidase N (1 μg/ml) and 25 μl of pancreatic carboxypeptidase A (1 μg/ml) were included in the same assay as controls. For the inhibition of carboxypeptidase B-like activity, the activated polypeptides were pre-incubated for 5 minutes with carboxypeptidase inhibitor from potato (1 μM final), and the residual activity was determined by adding TAFI Developer.

Carboxypeptidase activity of recombinant wild type TAFI and the polypeptides of the invention was also measured using other substrates with various C-terminal amino acid; hippuryl-arginine, hippuryl-lysine, hippuryl-leucine, and hippuryl-phenylalanine. Enzymes were incubated with substrates for 30 minute at room temperature and the rate of hydrolysis were determined by conversion of the product, hippuric acid, to a chromogen with cyanuric chloride dissolved in dioxane (Sigma). Briefly, in a 96-well microtiter plate, 12 μl of either activated wild type TAFI, the polypeptides of the invention, carboxypeptidase N or carboxypeptidase A (1 μg/ml each) was mixed with 24 μl of HEPES buffer (50 mM, pH 7.8), 12 μl H_2O and 12 μl of 25 mM substrate dissolved in 20 mM NaOH. The reaction was stopped by addition of 80 μl of phosphate buffer (0.2 M, pH 8.3) and 60 μl of 3% cyanuric chloride in

dioxane (w/v). After mixing thoroughly by pipetting, the clear supernatant was transferred to a new well, and absorbance was measured at 382 nm.

Wild type TAFI and the mutant TAFI polypeptides were activated by the thrombin-thrombomodulin complex at room temperature. At various time points aliquots were removed and analyzed by SDS-PAGE and Western blot. The time course of proteolytic activation of wild type TAFI and the polypeptides of the invention was similar when followed from 5 min to 20 min. The profiles after 10 min of activation are shown in Fig. 2. The zymogen with a molecular weight of 58 kDa was cleaved to yield activated polypeptides with a molecular weight of 35 kDa. Activation of the polypeptides of the invention was also calcium-dependent as shown for wild type TAFI (data not shown). Activation was more efficient when 150 mM NaCl was used in the activation buffer rather than 300 mM NaCl (data not shown). While all of the TAFI polypeptide mutants could be activated by thrombin-thrombomodulin complex, D256Q and D256Q, S207D were activated faster than wild type TAFI.

Carboxypeptidase B-like activity of activated wild type TAFI and the TAFI polypeptide mutants was compared by using a specific substrate for TAFIa (Fig.3). All of the mutant polypeptide lost most of the activity (they maintained only 2% to 6% of wild type TAFIa activity). Similar amounts of CPN and CPA, when tested in the same assay, had 15% and 2% of carboxypeptidase B-like activity (Fig.3).

The inhibition of wild type TAFI and the TAFI polypeptide mutants by CPI after activation with thrombin and thrombomodulin was measured. The percentage inhibition of TAFIa was calculated for each CPI concentration (Fig.4).

Substrates hippuryl-arginine, hippuryl-lysine, hippuryl-leucine, and hippuryl-phenylalanine were also used to measure carboxypeptidase activities of recombinant wild type TAFI and the mutant TAFI polypeptides, as well as similar amount of CPN and CPA as described above. The enzyme activity of each polypeptide is shown in (Fig.5).

EXAMPLE IV: Plasma Clot Lysis Assay

In a 96-well microtiter plate, 30 μl of human TAFI deficient plasma (Enzyme Research Lab), 12 μl of TAFI or the polypeptides of the invention were mixed with 3μl of 0.8 μM thrombomodulin, 5 μl of 24 μM CPI and 60 μl of assay buffer (40 mM HEPES, 150 mM NaCl, 0.02% Tween 80). The mixture was immediately added to another well containing 4 μl of 75 NIH units/ml thrombin, 2 μl of 1M $CaCl_2$ and 4 μl of 2 μg/ml Activase in separate aliquots. The total volume of the mixture was 120 μl. The final concentration of wild type TAFI or a TAFI polypeptide mutant in the solution was 0.2 μg/ml. After mixing by pipetting, the clot formation and lysis was monitored at 405 nm every 1 min at 37°C using a SpectraMAX 250 Microplate Spectrophotometer (Molecular Devices Corporation, Sunnyvale, CA). Lysis time was defined as the time at which turbidity is one-half the difference between the plateau reached after clotting and the base line value achieved at complete lysis.

The effect of wild type TAFI or the TAFI polypeptide mutants on fibrinolysis was determined in the plasma clot lysis assay as described above. Clotting induced by thrombin and subsequent fibrinolysis induced by tissue plasminogen activator were followed over time by measuring turbidity at 405 nm. The functional TAFI activity of wild type TAFI led to an increased plasma clot lysis time from 14 min. to 45 min. All of the mutant TAFI polypeptide had plasma clot lysis times of ~17 min, demonstrating that they possess much less functional TAFI activity than wild type TAFI. The addition of CPI, which inhibits TAFI activity, decreased the clot lysis time to 14 min in plasma with both wild type TAFI and with the TAFI polypeptide mutants (Fig.6A). This value is identical to that obtained in the absence of TAFI (Fig.6B), indicating that the extension of lysis time in these experiments is due to the carboypeptidase B activity derived from activated TAFI.

From the foregoing description, one skilled in the art can easily ascertain the essential characteristics of this invention, and without departing from the spirit and scope thereof, can make changes and modifications of the invention to adapt it to various usage and conditions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The preceding preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The entire disclosure of all applications, patents and publications, cited above and in the figures are hereby incorporated in their entirety by reference.

WHAT IS CLAIMED IS:

1. An isolated TAFI polypeptide mutant or a fragment or variant thereof, wherein the amino acid sequence of said polypeptide mutant comprises the sequence of wild-type TAFI (SEQ ID NO: 2), and wherein said wild-type sequence has been modified by the substitution of one or more amino acid residues involved in substrate binding.

2. The TAFI polypeptide mutant of Claim 1, wherein said mutations are at amino acid residues 207, 248, 256, or 257, wherein the numbering of said amino acid residues is taken from the amino acid sequence of activated TAFI (i.e. TAFIa; SEQ ID NO: 3).

3. The TAFI polypeptide mutant of Claim 2, wherein the Ser at amino acid position 207 in SEQ ID NO: 3 is replaced with Asp, e.g. S207D.

4. The TAFI polypeptide mutant of Claim 2, wherein the Leu at amino acid position 248 in SEQ ID NO: 3 is replaced with Try, e.g. L248W.

5. The TAFI polypeptide mutant of Claim 2, wherein the Asp at amino acid position 256 in SEQ ID NO: 3 is replaced with Glu, e.g. D256Q.

6. The TAFI polypeptide mutant of Claim 2, wherein the Asp at amino acid position 257 in SEQ ID NO: 3 is replaced with Ala., e.g. D257A.

7. The TAFI polypeptide mutant of Claim 2, wherein the Ser at amino acid position 207 in SEQ ID NO: 3 is replaced with Asp, e.g. S207D, and the Asp at amino acid position 256 in SEQ ID NO: 3 is replaced with Glu., e.g. D256Q.

8. The TAFI polypeptide mutant of Claim 2, wherein said polypeptide further comprises an additional mutation at an amino acid position 69, 125,143,197,198, 199, 249, or 271 of SEQ ID NO: 3.

9. An isolated polynucleotide encoding the polypeptide mutant of Claim 1.

10. A recombinant vector comprising the polynucleotide of Claim 9 operably linked to at least one regulatory sequence.

11. The recombinant vector of Claim 10, wherein said regulatory sequence is a promoter.

12. The recombinant vector of Claim 10, wherein said vector is a baculovirus expression vector or a mammalian vector.

13. A cell comprising the vector of Claim 10.

14. The cell of Claim 13, wherein said cell is a mammalian, a yeast, or an insect cell.

15. The cell of Claim 14, wherein said cell is an insect cell.

16. A method of making a polypeptide of Claim 1 comprising transforming a cell with said vector of Claim 10 , incubating the cell under conditions in which the polypeptide is expressed, and harvesting the polypeptide.

17. A method of screening for agents that modulate the activity of said TAFI polypeptide mutant of Claim 1 or a related carboxypeptidase, comprising contacting a cell expressing said mutant *or a tissue cell extract* with a putative modulatory agent, and measuring the activity of said polypeptide mutant.

18. A method of screening for agents that modulate the expression or activity of said TAFI polypeptide mutant of Claim 1 or a related carboxypeptidase, comprising contacting a cell expressing said mutant or a tissue cell extract with a putative modulatory agent, and *monitoring plasma and/or tissue levels of said polypeptide or related* carboxypeptidase.

19. A method of screening for agents that modulate the expression or activity of a polynucleotide which encodes said TAFI polypeptide mutant of Claim 1 or a related carboxypeptidase, comprising contacting a cell transformed with said polynucleotide *or a tissue cell extract* with a putative modulatory agent, and measuring the amount or activity of said polynucleotide.

20. The method of Claim 19, wherein said agent reduces the activity of said polynucleotide..

21. The method of Claim 19, wherein said agent enhances the activity of said polynucleotide.

22. A method for altering the binding specificity of a substrate of a polypeptide of SEQ ID NO: 1 or a related carboxypeptidase, comprising mutating or altering the amino acids involved in substrate binding.

23. A method of screening for an agent which binds to said TAFI polypeptide mutant of Claim 1 or a related carboxypeptidase, comprising contacting said polypeptide or related carboxypeptidase with a

putative binding agent and determining the presence of a bound complex.

24. The method of claim 23, wherein said agent is a substrate.

5

FIGURE 1

```
TAFI  MKLCSLAVLV PIVLFCEQHV FAFQSGQVLA ALPRTSRQVQ VLQNLTTTYE
CPN   .......... .......... .......... .......... ..........

TAFI  IVLWQPVTAD LIVKKKQVHF FVNASDVDNV KAHLNVSGIP CSVLLADVED
CPN   .......... .......... .......... .......... .......MSD

                        1          10         20         30
TAFI  LIQQQISNDT VSPRASASYY EQYHSLNEIY SWIEFITERH PDMLTKIHIG
CPN   LLSVFL.HLL LLFKLVAPVT FRHHRYDDLV RTLYKVQNEC PGITRVYSIG
                         ^
         40         50                60         70         80
TAFI  SSFEKYPLYV LKVSGKEQTA KN...AIWID CGIHAREWIS PAFCLWFIGH
CPN   RSVEGRHLYV LEFSDHPGIH EPLEPEVKYV GNMHGNEALG RELMLQLSEF
                                          * #*
              90         100        110        120        130
TAFI  ITQFYG.IIG QYTNLLRLVD FYVMPVVNVD GYDYSWKKNR MWRKNRSFYA
CPN   LCEEFRNRNQ RIVQLIQDTR IHILPSMNPD GYEVAAA..Q GPNKPGYLVG
                                                   &/#
              140             150        160         170
TAFI  NNHCIGTDLN RNFASKH... WCEEGASSSS CSETYCGLY. PESEPEVKAV
CPN   RNNANGVDLN RNFPDLNTYI YYNEKYGGPN HHLPLPDNWK SQVEPETRAV
                 # #
       180        190         200        210                 220
TAFI  ASFLRRNINQ IKAYISMHSY SQHIVFPYSY TRSK.SKDHE EL.......S
CPN   IRWMHSFNFV LSA..NLHGG AVVANYPYDK SFEHRVRGVR RTASTPTPDD
                         *## #       #
                230        240                 250        260
TAFI  LVASEAVRAI EKTSKNTRYT .......... HGHGSETLYL APGGGDDWIY
CPN   KLFQKLAKVY SYAH.GWMFQ GWNCGDYFPD GITNGASWYS LSKGMQDFNY
                                                 #        #
                270        280        290        300
TAFI  DLGIKYSFTI ELRDTGTYGF LLPERYIKPT CREAFAAVSK IAWHVIRNV.
CPN   LHTNCFEITL EL....SCDK FPPEEELQRE WLGNREALIQ FLEQVHQGIK
                  &/#

TAFI  .......... .......... .......... .......... ..........
CPN   GMVLDENYNN LANAVISVSG INHDVTSGDH GDYFRLLLPG IYTVSATAPG

TAFI  .......... .......... .......... .......... ..........
CPN   YDPETVTVTV GPAEPTLVNF HLKRSIPQVS PVRRAPSRRH GV.RAKVQPQ

TAFI  .......... ......
CPN   ARKKEMEMRQ LQRGPA
```

Numbering starts with alanine of activated TAFI
^ = cleavage site
*=zinc binding; #=known or potential substrate binding; &=catalytic activity

FIGURE 2

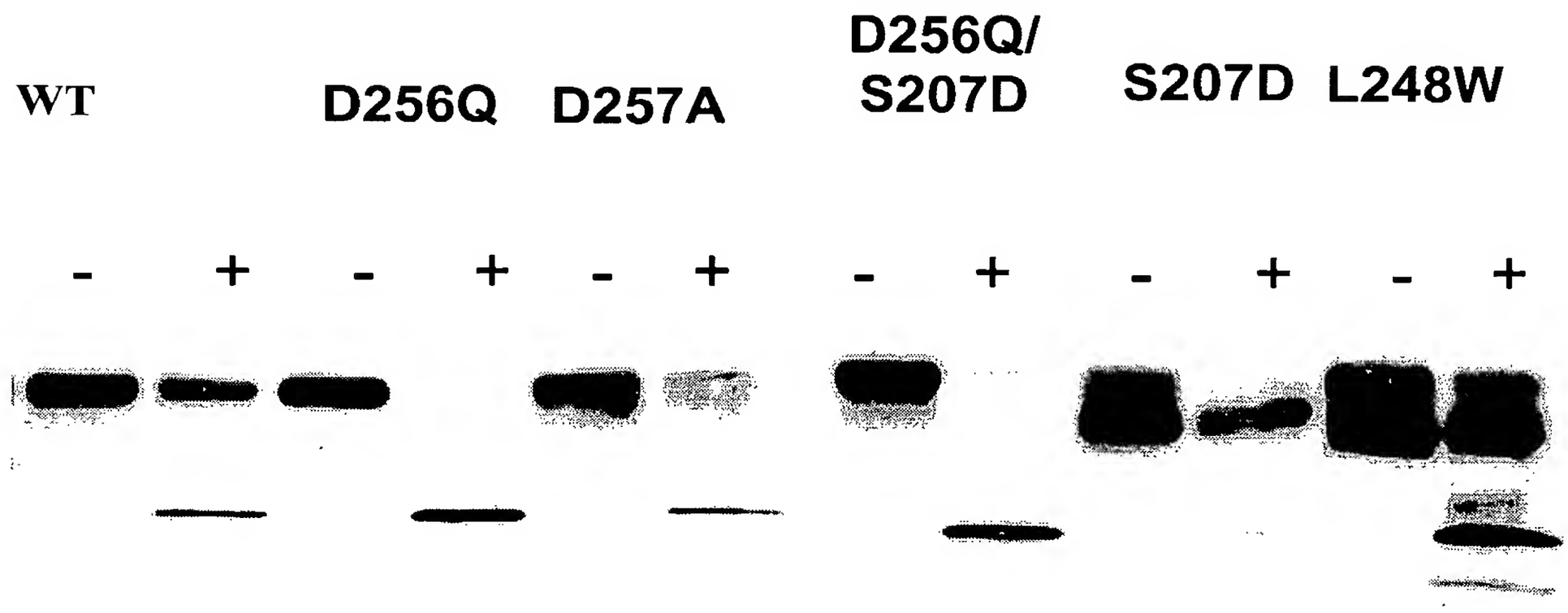


FIGURE 3

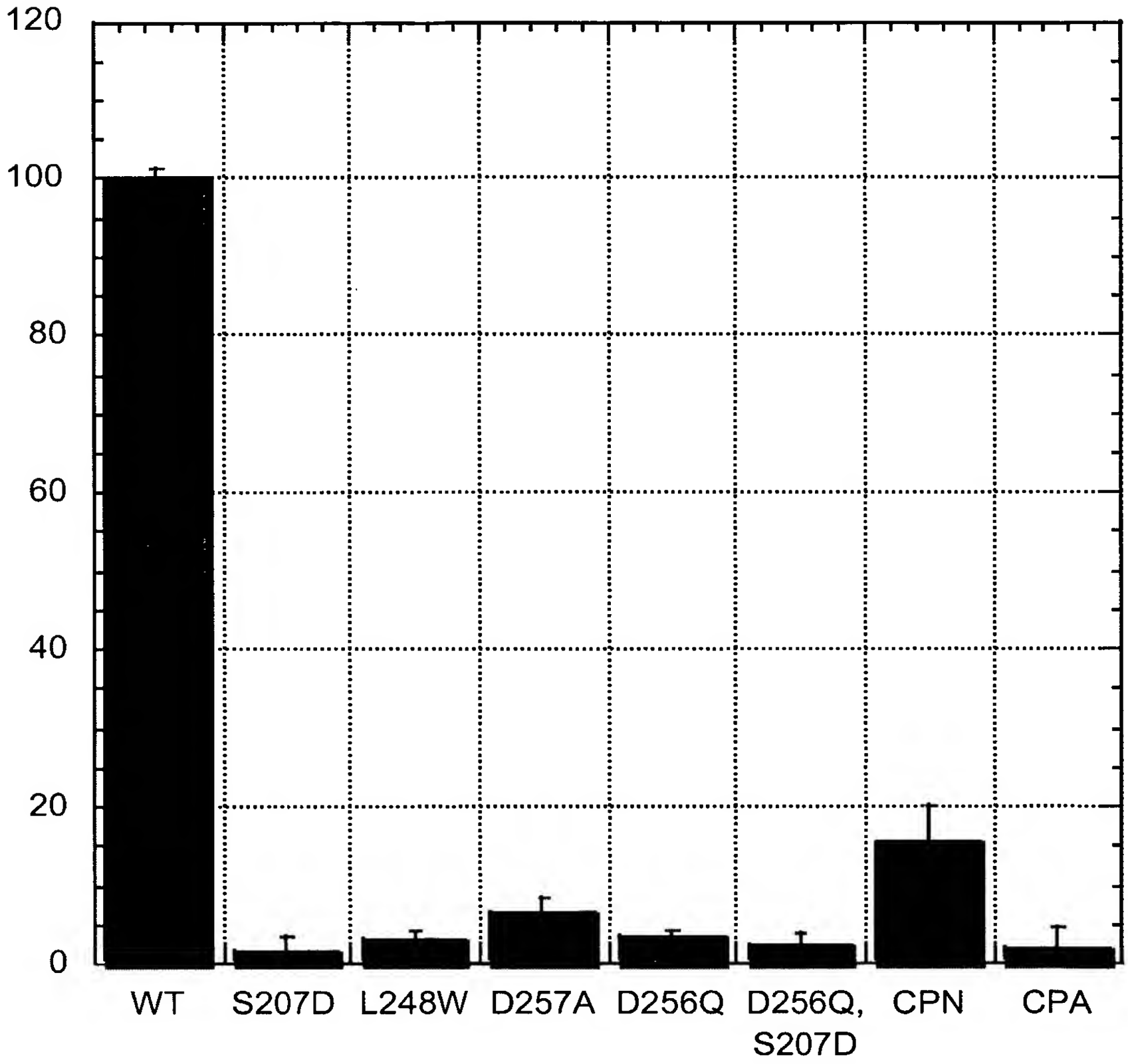

120
100
80
60
40
20
0
% of WT TAFIa activity
WT
S207D
L248W
D257A
D256Q
D256Q, S207D
CPN
CPA

FIGURE 4

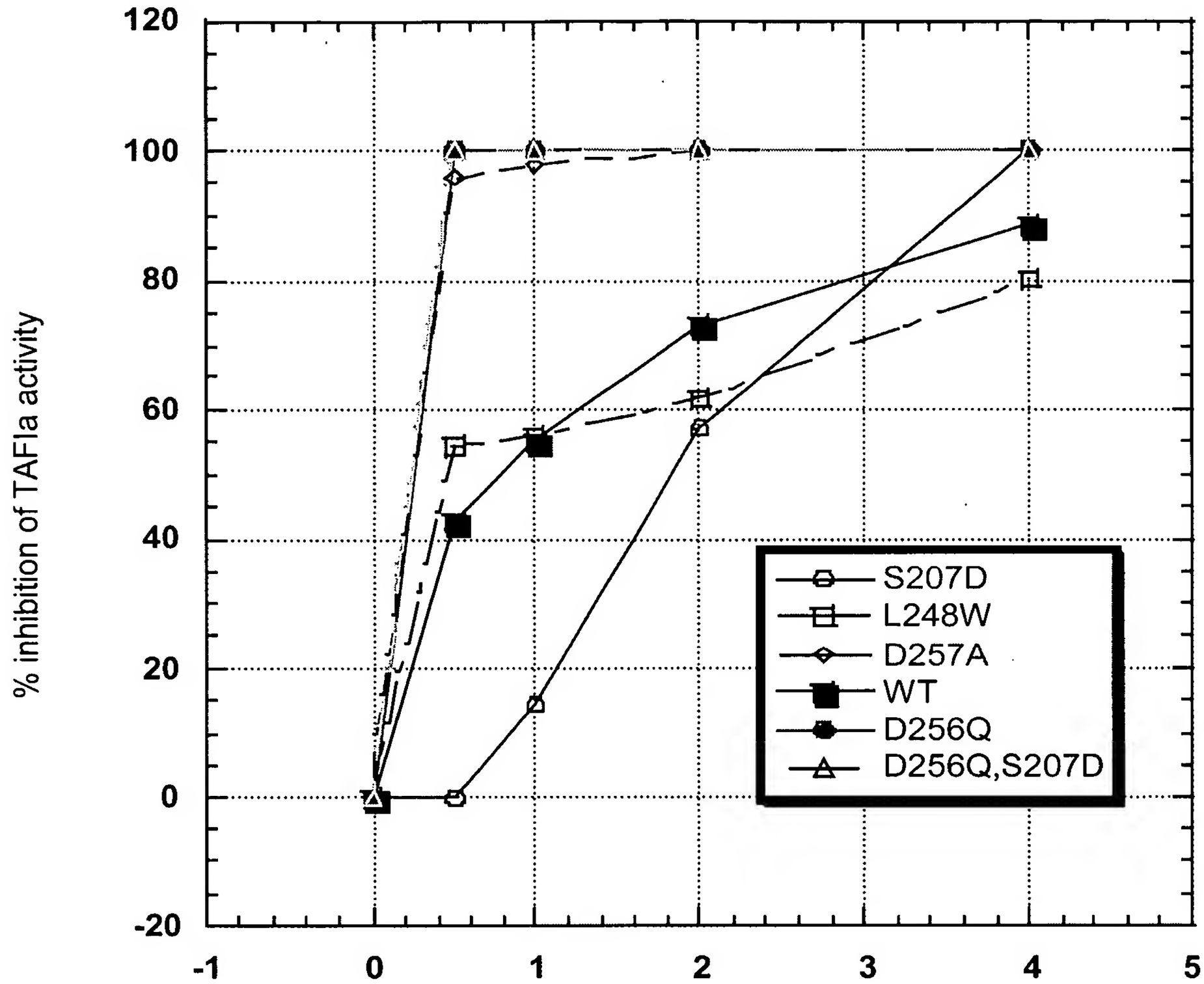

120
100
80
60
40
20
0
-20
-1
0
1
2
3
4
5
% inhibition of TAFIa activity
S207D
L248W
D257A
WT
D256Q
D256Q,S207D
CPI concentration (uM)

FIGURE 5

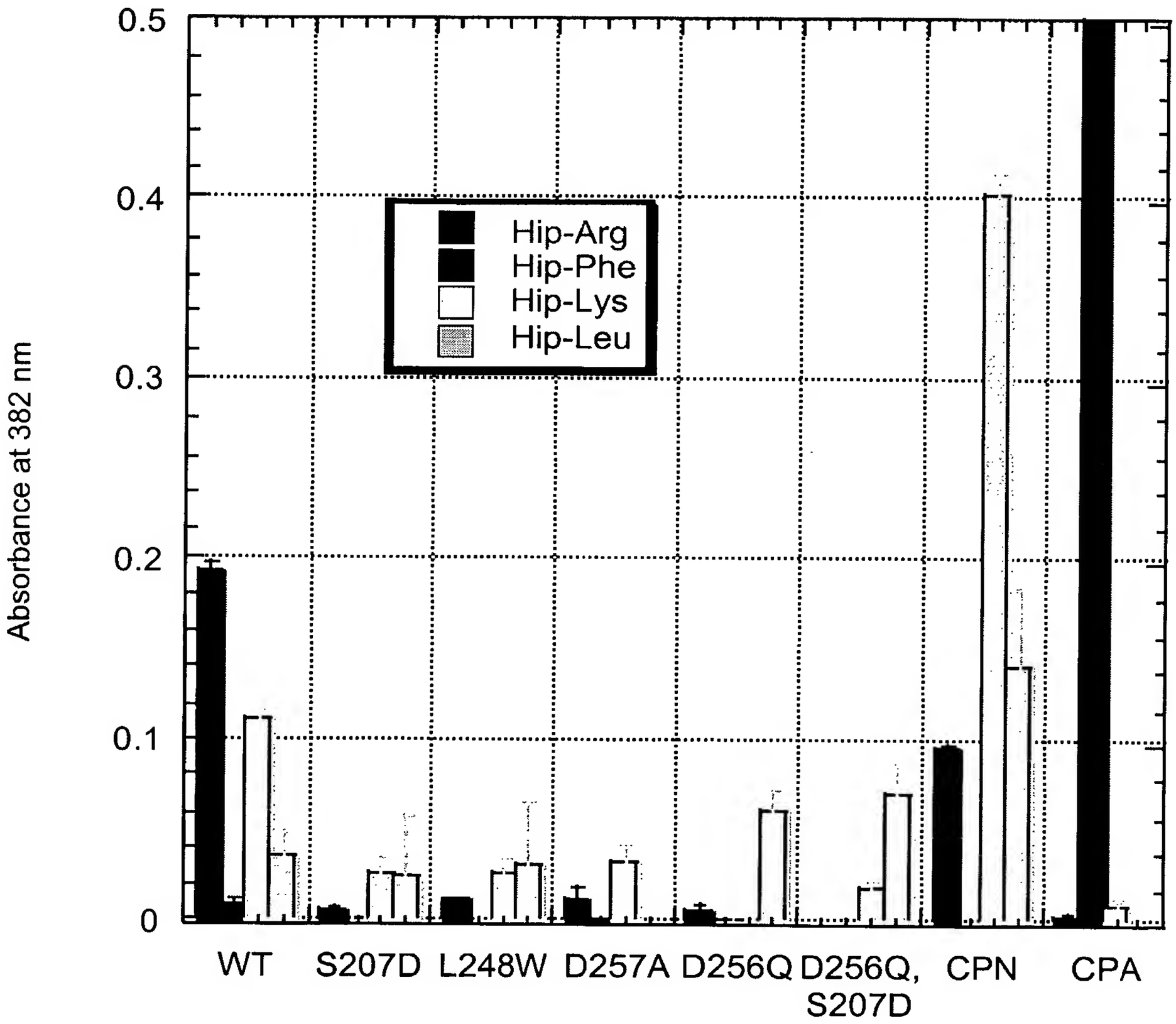

Absorbance at 382 nm
0.5
0.4
0.3
0.2
0.1
0
Hip-Arg
Hip-Phe
Hip-Lys
Hip-Leu
WT
S207D
L248W
D257A
D256Q
D256Q, S207D
CPN
CPA

FIGURE 6A

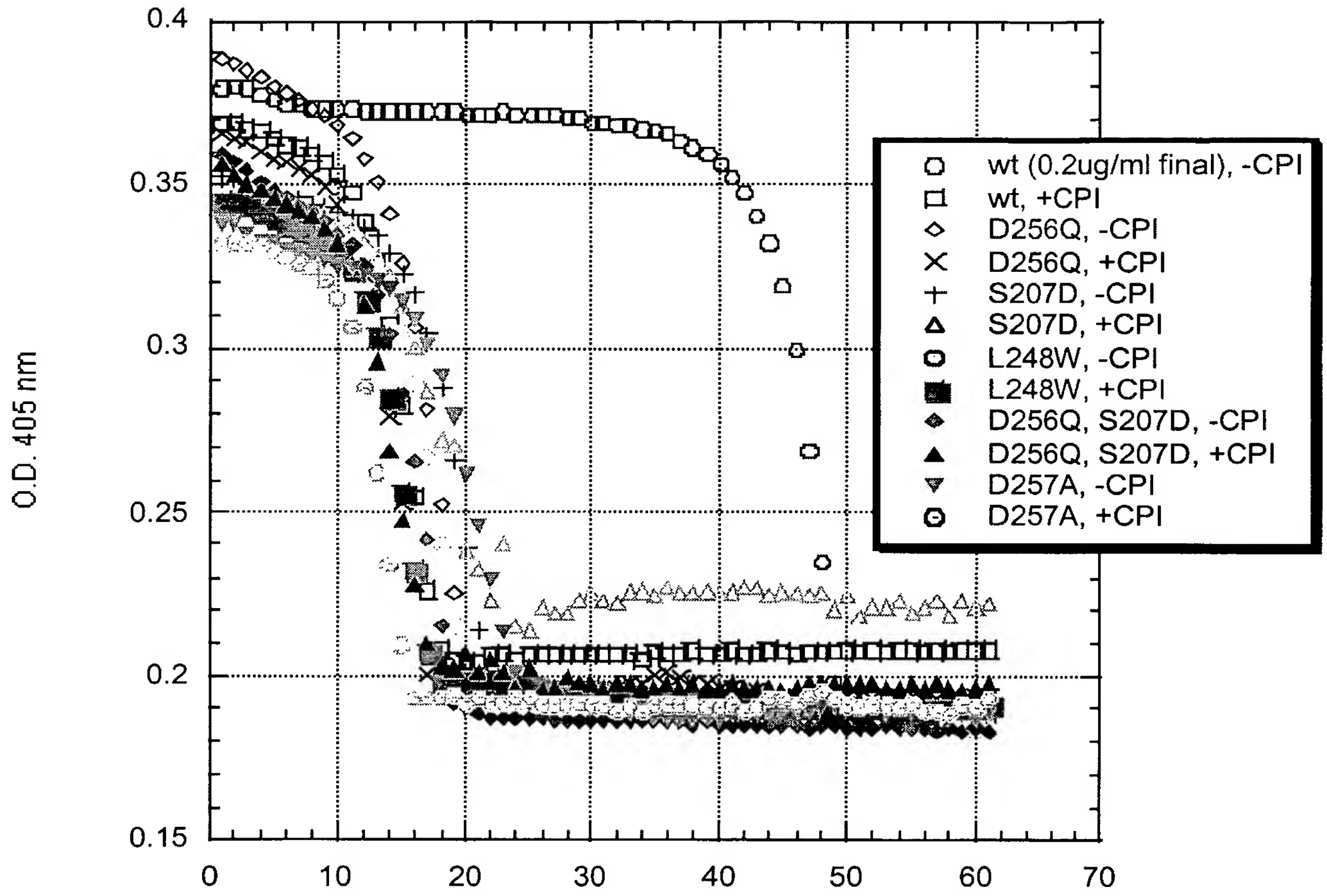

0.4
0.35
0.3
0.25
0.2
0.15
O.D. 405 nm
0
10
20
30
40
50
60
70
min
wt (0.2ug/ml final), -CPI
wt, +CPI
D256Q, -CPI
D256Q, +CPI
S207D, -CPI
S207D, +CPI
L248W, -CPI
L248W, +CPI
D256Q, S207D, -CPI
D256Q, S207D, +CPI
D257A, -CPI
D257A, +CPI

FIGURE 6B

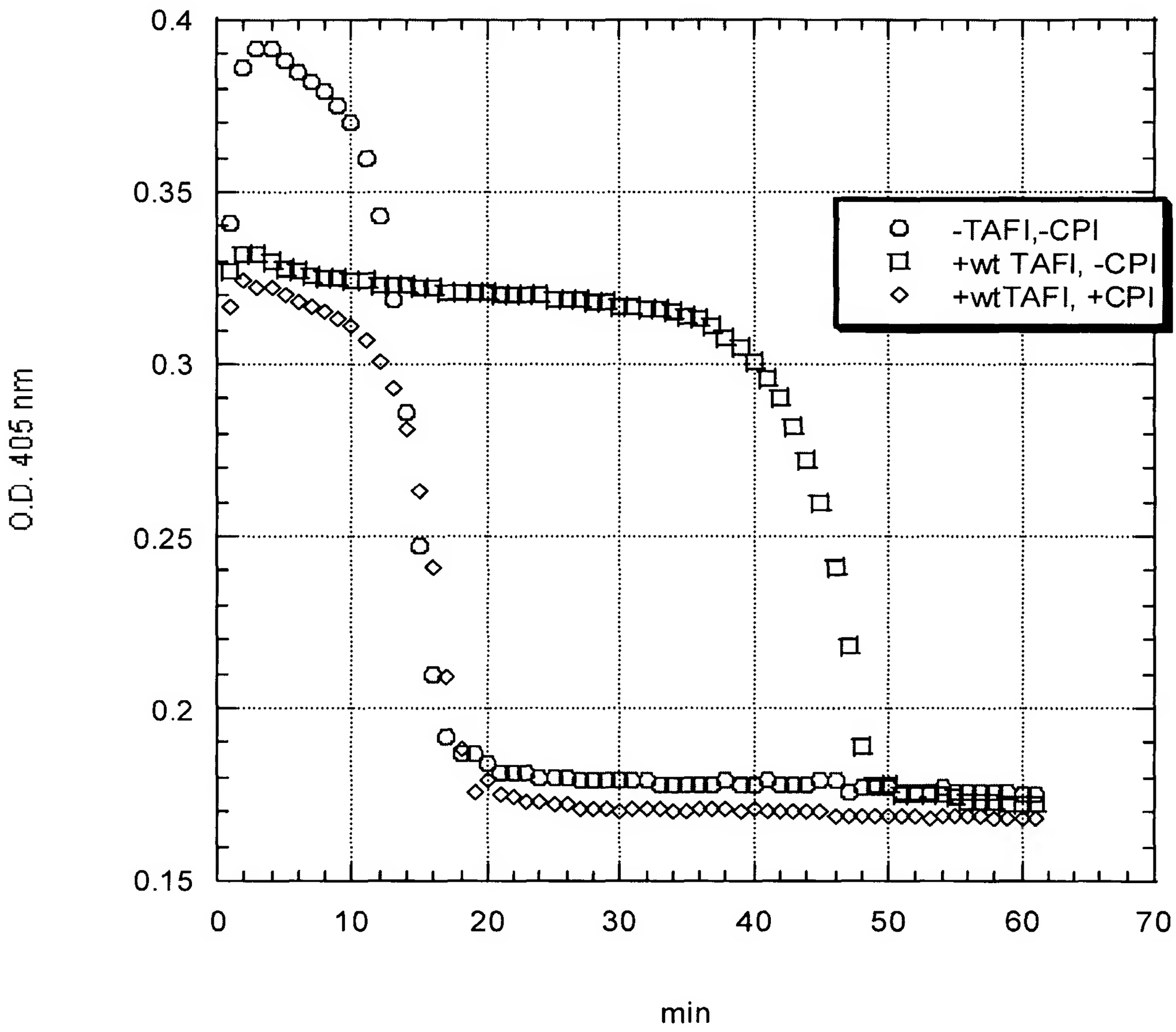
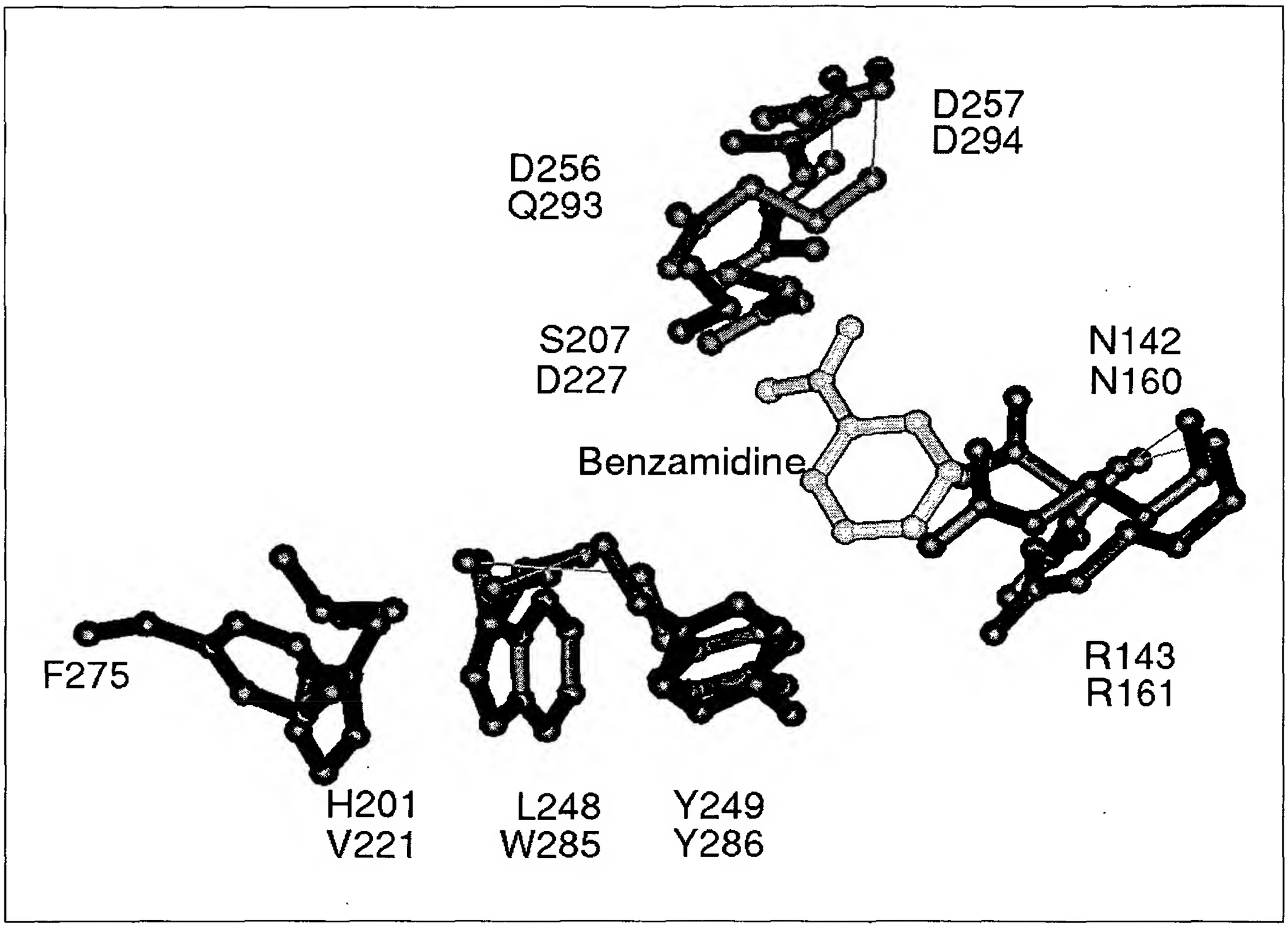

0.4
0.35
0.3
0.25
0.2
0.15
O.D. 405 nm
0
10
20
30
40
50
60
70
min
-TAFI,-CPI
+wt TAFI, -CPI
+wtTAFI, +CPI

FIGURE 8


D257
D294
D256
Q293
S207
D227
N142
N160
Benzamidine
R143
R161
F275
H201
V221
L248
W285
Y249
Y286

SEQUENCE LISTING

<110> Schering Aktiengesellschaft
McCarrick, Margaret
Morser, Michael J
Nagashima, Mariko
Zhao, Lei

<120> Carboxypeptidase B Related Polypeptides and Methods of Use

<130> 53029AWOM1

<150> US 60/406,341
<151> 2002-08-28

<160> 11

<170> PatentIn version 3.2

<210> 1
<211> 1272
<212> DNA
<213> *Homo sapiens*

<220>
<221> CDS
<222> (1)..(1272)

<220>
<221> sig_peptide
<222> (1)..(66)

<400> 1

```
atg aag ctt tgc agc ctt gca gtc ctt gta ccc att gtt ctc ttc tgt        48
Met Lys Leu Cys Ser Leu Ala Val Leu Val Pro Ile Val Leu Phe Cys
1               5                  10                  15

gag cag cat gtc ttc gcg ttt cag agt ggc caa gtt cta gct gct ctt        96
Glu Gln His Val Phe Ala Phe Gln Ser Gly Gln Val Leu Ala Ala Leu
            20                  25                  30

cct aga acc tct agg caa gtt caa gtt cta cag aat ctt act aca aca       144
Pro Arg Thr Ser Arg Gln Val Gln Val Leu Gln Asn Leu Thr Thr Thr
        35                  40                  45

tat gag att gtt ctc tgg cag ccg gta aca gct gac ctt att gtg aag       192
Tyr Glu Ile Val Leu Trp Gln Pro Val Thr Ala Asp Leu Ile Val Lys
    50                  55                  60

aaa aaa caa gtc cat ttt ttt gta aat gca tct gat gtc gac aat gtg       240
Lys Lys Gln Val His Phe Phe Val Asn Ala Ser Asp Val Asp Asn Val
65                  70                  75                  80

aaa gcc cat tta aat gtg agc gga att cca tgc agt gtc ttg ctg gca       288
Lys Ala His Leu Asn Val Ser Gly Ile Pro Cys Ser Val Leu Leu Ala
                85                  90                  95

gac gtg gaa gat ctt att caa cag cag att tcc aac gac aca gtc agc       336
Asp Val Glu Asp Leu Ile Gln Gln Gln Ile Ser Asn Asp Thr Val Ser
            100                 105                 110

ccc cga gcc tcc gca tcg tac tat gaa cag tat cac tca cta aat gaa       384
Pro Arg Ala Ser Ala Ser Tyr Tyr Glu Gln Tyr His Ser Leu Asn Glu
        115                 120                 125
```

```
atc tat tct tgg ata gaa ttt ata act gag agg cat cct gat atg ctt      432
Ile Tyr Ser Trp Ile Glu Phe Ile Thr Glu Arg His Pro Asp Met Leu
    130                 135                 140

aca aaa atc cac att gga tcc tca ttt gag aag tac cca ctc tat gtt      480
Thr Lys Ile His Ile Gly Ser Ser Phe Glu Lys Tyr Pro Leu Tyr Val
145                 150                 155                 160

tta aag gtt tct gga aaa gaa caa aca gcc aaa aat gcc ata tgg att      528
Leu Lys Val Ser Gly Lys Glu Gln Thr Ala Lys Asn Ala Ile Trp Ile
                165                 170                 175

gac tgt gga atc cat gcc aga gaa tgg atc tct cct gct ttc tgc ttg      576
Asp Cys Gly Ile His Ala Arg Glu Trp Ile Ser Pro Ala Phe Cys Leu
            180                 185                 190

tgg ttc ata ggc cat ata act caa ttc tat ggg ata ata ggg caa tat      624
Trp Phe Ile Gly His Ile Thr Gln Phe Tyr Gly Ile Ile Gly Gln Tyr
        195                 200                 205

acc aat ctc ctg agg ctt gtg gat ttc tat gtt atg ccg gtg gtt aat      672
Thr Asn Leu Leu Arg Leu Val Asp Phe Tyr Val Met Pro Val Val Asn
    210                 215                 220

gtg gac ggt tat gac tac tca tgg aaa aag aat cga atg tgg aga aag      720
Val Asp Gly Tyr Asp Tyr Ser Trp Lys Lys Asn Arg Met Trp Arg Lys
225                 230                 235                 240

aac cgt tct ttc tat gcg aac aat cat tgc atc gga aca gac ctg aat      768
Asn Arg Ser Phe Tyr Ala Asn Asn His Cys Ile Gly Thr Asp Leu Asn
                245                 250                 255

agg aac ttt gct tcc aaa cac tgg tgt gag gaa ggt gca tcc agt tcc      816
Arg Asn Phe Ala Ser Lys His Trp Cys Glu Glu Gly Ala Ser Ser Ser
            260                 265                 270

tca tgc tcg gaa acc tac tgt gga ctt tat cct gag tca gaa cca gaa      864
Ser Cys Ser Glu Thr Tyr Cys Gly Leu Tyr Pro Glu Ser Glu Pro Glu
        275                 280                 285

gtg aag gca gtg gct agt ttc ttg aga aga aat atc aac cag att aaa      912
Val Lys Ala Val Ala Ser Phe Leu Arg Arg Asn Ile Asn Gln Ile Lys
    290                 295                 300

gca tac atc agc atg cat tca tac tcc cag cat ata gtg ttt cca tat      960
Ala Tyr Ile Ser Met His Ser Tyr Ser Gln His Ile Val Phe Pro Tyr
305                 310                 315                 320

tcc tat aca cga agt aaa agc aaa gac cat gag gaa ctg tct cta gta     1008
Ser Tyr Thr Arg Ser Lys Ser Lys Asp His Glu Glu Leu Ser Leu Val
                325                 330                 335

gcc agt gaa gca gtt cgt gct att gag aaa act agt aaa aat acc agg     1056
Ala Ser Glu Ala Val Arg Ala Ile Glu Lys Thr Ser Lys Asn Thr Arg
            340                 345                 350

tat aca cat ggc cat ggc tca gaa acc tta tac cta gct cct gga ggt     1104
Tyr Thr His Gly His Gly Ser Glu Thr Leu Tyr Leu Ala Pro Gly Gly
        355                 360                 365

ggg gac gat tgg atc tat gat ttg ggc atc aaa tat tcg ttt aca att     1152
Gly Asp Asp Trp Ile Tyr Asp Leu Gly Ile Lys Tyr Ser Phe Thr Ile
    370                 375                 380
```

```
gaa ctt cga gat acg ggc aca tac gga ttc ttg ctg ccg gag cgt tac      1200
Glu Leu Arg Asp Thr Gly Thr Tyr Gly Phe Leu Leu Pro Glu Arg Tyr
385                 390                 395                 400

atc aaa ccc acc tgt aga gaa gct ttt gcc gct gtc tct aaa ata gct      1248
Ile Lys Pro Thr Cys Arg Glu Ala Phe Ala Ala Val Ser Lys Ile Ala
                405                 410                 415

tgg cat gtc att agg aat gtt taa                                     1272
Trp His Val Ile Arg Asn Val
            420
```

<210> 2
<211> 423
<212> PRT
<213> Homo sapiens

<400> 2

```
Met Lys Leu Cys Ser Leu Ala Val Leu Val Pro Ile Val Leu Phe Cys
1               5                   10                  15

Glu Gln His Val Phe Ala Phe Gln Ser Gly Gln Val Leu Ala Ala Leu
            20                  25                  30

Pro Arg Thr Ser Arg Gln Val Gln Val Leu Gln Asn Leu Thr Thr Thr
        35                  40                  45

Tyr Glu Ile Val Leu Trp Gln Pro Val Thr Ala Asp Leu Ile Val Lys
    50                  55                  60

Lys Lys Gln Val His Phe Phe Val Asn Ala Ser Asp Val Asp Asn Val
65                  70                  75                  80

Lys Ala His Leu Asn Val Ser Gly Ile Pro Cys Ser Val Leu Leu Ala
                85                  90                  95

Asp Val Glu Asp Leu Ile Gln Gln Gln Ile Ser Asn Asp Thr Val Ser
            100                 105                 110

Pro Arg Ala Ser Ala Ser Tyr Tyr Glu Gln Tyr His Ser Leu Asn Glu
        115                 120                 125

Ile Tyr Ser Trp Ile Glu Phe Ile Thr Glu Arg His Pro Asp Met Leu
    130                 135                 140

Thr Lys Ile His Ile Gly Ser Ser Phe Glu Lys Tyr Pro Leu Tyr Val
145                 150                 155                 160

Leu Lys Val Ser Gly Lys Glu Gln Thr Ala Lys Asn Ala Ile Trp Ile
                165                 170                 175

Asp Cys Gly Ile His Ala Arg Glu Trp Ile Ser Pro Ala Phe Cys Leu
```

180 185 190

Trp Phe Ile Gly His Ile Thr Gln Phe Tyr Gly Ile Ile Gly Gln Tyr
195 200 205

Thr Asn Leu Leu Arg Leu Val Asp Phe Tyr Val Met Pro Val Val Asn
210 215 220

Val Asp Gly Tyr Asp Tyr Ser Trp Lys Lys Asn Arg Met Trp Arg Lys
225 230 235 240

Asn Arg Ser Phe Tyr Ala Asn Asn His Cys Ile Gly Thr Asp Leu Asn
245 250 255

Arg Asn Phe Ala Ser Lys His Trp Cys Glu Glu Gly Ala Ser Ser Ser
260 265 270

Ser Cys Ser Glu Thr Tyr Cys Gly Leu Tyr Pro Glu Ser Glu Pro Glu
275 280 285

Val Lys Ala Val Ala Ser Phe Leu Arg Arg Asn Ile Asn Gln Ile Lys
290 295 300

Ala Tyr Ile Ser Met His Ser Tyr Ser Gln His Ile Val Phe Pro Tyr
305 310 315 320

Ser Tyr Thr Arg Ser Lys Ser Lys Asp His Glu Glu Leu Ser Leu Val
325 330 335

Ala Ser Glu Ala Val Arg Ala Ile Glu Lys Thr Ser Lys Asn Thr Arg
340 345 350

Tyr Thr His Gly His Gly Ser Glu Thr Leu Tyr Leu Ala Pro Gly Gly
355 360 365

Gly Asp Asp Trp Ile Tyr Asp Leu Gly Ile Lys Tyr Ser Phe Thr Ile
370 375 380

Glu Leu Arg Asp Thr Gly Thr Tyr Gly Phe Leu Leu Pro Glu Arg Tyr
385 390 395 400

Ile Lys Pro Thr Cys Arg Glu Ala Phe Ala Ala Val Ser Lys Ile Ala
405 410 415

Trp His Val Ile Arg Asn Val
420

<210> 3
<211> 309

<212> PRT
<213> Homo sapiens

<400> 3

Ala Ser Ala Ser Tyr Tyr Glu Gln Tyr His Ser Leu Asn Glu Ile Tyr
1 5 10 15

Ser Trp Ile Glu Phe Ile Thr Glu Arg His Pro Asp Met Leu Thr Lys
20 25 30

Ile His Ile Gly Ser Ser Phe Glu Lys Tyr Pro Leu Tyr Val Leu Lys
35 40 45

Val Ser Gly Lys Glu Gln Thr Ala Lys Asn Ala Ile Trp Ile Asp Cys
50 55 60

Gly Ile His Ala Arg Glu Trp Ile Ser Pro Ala Phe Cys Leu Trp Phe
65 70 75 80

Ile Gly His Ile Thr Gln Phe Tyr Gly Ile Ile Gly Gln Tyr Thr Asn
85 90 95

Leu Leu Arg Leu Val Asp Phe Tyr Val Met Pro Val Val Asn Val Asp
100 105 110

Gly Tyr Asp Tyr Ser Trp Lys Lys Asn Arg Met Trp Arg Lys Asn Arg
115 120 125

Ser Phe Tyr Ala Asn Asn His Cys Ile Gly Thr Asp Leu Asn Arg Asn
130 135 140

Phe Ala Ser Lys His Trp Cys Glu Glu Gly Ala Ser Ser Ser Ser Cys
145 150 155 160

Ser Glu Thr Tyr Cys Gly Leu Tyr Pro Glu Ser Glu Pro Glu Val Lys
165 170 175

Ala Val Ala Ser Phe Leu Arg Arg Asn Ile Asn Gln Ile Lys Ala Tyr
180 185 190

Ile Ser Met His Ser Tyr Ser Gln His Ile Val Phe Pro Tyr Ser Tyr
195 200 205

Thr Arg Ser Lys Ser Lys Asp His Glu Glu Leu Ser Leu Val Ala Ser
210 215 220

Glu Ala Val Arg Ala Ile Glu Lys Thr Ser Lys Asn Thr Arg Tyr Thr
225 230 235 240

His Gly His Gly Ser Glu Thr Leu Tyr Leu Ala Pro Gly Gly Gly Asp
245 250 255

Asp Trp Ile Tyr Asp Leu Gly Ile Lys Tyr Ser Phe Thr Ile Glu Leu
260 265 270

Arg Asp Thr Gly Thr Tyr Gly Phe Leu Leu Pro Glu Arg Tyr Ile Lys
275 280 285

Pro Thr Cys Arg Glu Ala Phe Ala Ala Val Ser Lys Ile Ala Trp His
290 295 300

Val Ile Arg Asn Val
305

<210> 4
<211> 27
<212> DNA
<213> artificial

<220>
<223> D256primer1

<400> 4
atagatccaa tcctgcccac ctccagg 27

<210> 5
<211> 27
<212> DNA
<213> artificial

<220>
<223> D256primer2

<400> 5
cctggaggtg ggcaggattg gatctat 27

<210> 6
<211> 37
<212> DNA
<213> artificial

<220>
<223> S207primer1

<400> 6
gggacgattg gatctatcta ttgggcatca aatattc 37

<210> 7
<211> 37
<212> DNA
<213> artificial

<220>
<223> S207primer2

<400> 7

gaatatttga tgcccaatag atagatccaa tcgtccc 37

<210> 8
<211> 28
<212> DNA
<213> artificial

<220>
<223> L248primer1

<400> 8
ggctcagaaa cctggtacct agctcctg 28

<210> 9
<211> 28
<212> DNA
<213> artificial

<220>
<223> L248primer2

<400> 9
caggagctag gtaccaggtt tctgagcc 28

<210> 10
<211> 27
<212> DNA
<213> artificial

<220>
<223> D257primer1

<400> 10
tggaggtggg gatgcttgga tctatga 27

<210> 11
<211> 27
<212> DNA
<213> artificial

<220>
<223> D257primer2

<400> 11
tcatagatcc aagcatcccc acctcca 27

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
11 March 2004 (11.03.2004)

PCT

(10) International Publication Number
WO 2004/020587 A3

(51) International Patent Classification[7]: C12N 9/48, 1/20, 15/00, C07H 21/04

(21) International Application Number: PCT/US2003/026773

(22) International Filing Date: 25 August 2003 (25.08.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/406,341 28 August 2002 (28.08.2002) US

(71) Applicant *(for all designated States except US)*: **SCHERING AKTIENGESELLSCHAFT** [DE/DE]; Mueller-strasse 178, 13342 Berlin (DE).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **McCARRICK, Margaret, Anne** [US/US]; 3721 Balboa Terrace, Unit E, San Diego, CA 92117 (US). **MORSER, Michael, John** [GB/US]; 746 Sanchez Street, San Francisco, CA 94114 (US). **NAGASHIMA, Mariko** [US/US]; 2409 Palmer Avenue, Belmont, CA 94002 (CA). **ZHAO, Lei** [CN/US]; 21 Palmer Court, Pleasant Hill, CA 94523 (US).

(74) Agent: **WASHTIEN, Wendy, L.**; Berlex Biosciences, 2600 Hilltop Drive, P.O. Box 4099, Richmond, CA 94804 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

(88) Date of publication of the international search report:
13 January 2005

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 2004/020587 A3

(54) Title: CARBOXYPEPTIDASE B RELATED POLYPEPTIDES AND METHODS OF USE

(57) Abstract: Recombinant TAFI polypeptide mutants, as well as the polynucleotides encoding such polypeptides, are disclosed. Also disclosed are methods for utilizing such polypeptides in methods of screening for potential therapeutic agents, substrates, inhibitors or enhancers of carboxypeptidases.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US03/26773

A. CLASSIFICATION OF SUBJECT MATTER
IPC(7) : C12N 9/48, 1/20, 15/00; C07H 21/04
US CL : 435/212, 252.3, 320.1; 536/23.2
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 435/212, 252.3, 320.1; 536/23.2

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
WEST, STN AND SEQUENCE DATA BASE FOR SEQ ID NOS: 1-3.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98/55645 A1 (SCHERING AKTIENGESELLSCHAFT) 10 December 1998 (10.12.1998), see Examples 1-6, SEQ ID Nos. 1 (DNA) and 2 (protein) being 100% identical to Applicants' SEQ ID NO : 1 & 2.	1, 9-14, 16, 22
X	US 5,474,901 A (DRAYNA et al.) 12 December 1995 (12.12.1995), see abstarct, summary, Table 1, column 11 & 12 showing the mutations sites and regions including 207 though not with reference to Applicant's SEQ ID NO : 3. Applicants' SEQ ID NO : 2 is 100% identical to SEQ ID NO : 3 of the patent.	1, 9-16,
A	EATON et al. Isolation, Molecular Clonning, and Partial Characterization of a Novel CarboxypeptidaseB from Human Plasma. J. Biol. Chem., 15 November 1991, Vol. 266, No. 32, pages 21833-21838.	1-24

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
04 March 2004 (04.03.2004)	24 MAY 2004
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450 Alexandria, Virginia 22313-1450 Facsimile No. (703) 305-3230	Authorized officer Tekchand Saidha Telephone No. 571/272-1600

Form PCT/ISA/210 (second sheet) (July 1998)